

(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2005-505070

(P2005-505070A)

(43) 公表日 平成17年2月17日(2005.2.17)

(51) Int. Cl.⁷

G06F 17/30

F I

G06F 17/30 340B

テーマコード (参考)

5B075

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2003-533591 (P2003-533591)
 (86) (22) 出願日 平成14年9月10日 (2002. 9. 10)
 (85) 翻訳文提出日 平成16年3月29日 (2004. 3. 29)
 (86) 国際出願番号 PCT/IB2002/003693
 (87) 国際公開番号 W02003/030528
 (87) 国際公開日 平成15年4月10日 (2003. 4. 10)
 (31) 優先権主張番号 09/966, 414
 (32) 優先日 平成13年9月28日 (2001. 9. 28)
 (33) 優先権主張国 米国 (US)
 (81) 指定国 EP (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), CN, JP, KR

(71) 出願人 590000248
 コーニンクレッカ フィリップス エレクトロニクス エヌ ヴィ
 Koninklijke Philips Electronics N. V.
 オランダ国 5621 ペーアー アインドーフエン フルーネヴァウツウェeg 1
 Groenewoudseweg 1, 5621 BA Eindhoven, The Netherlands
 (74) 代理人 100092048
 弁理士 沢田 雅男

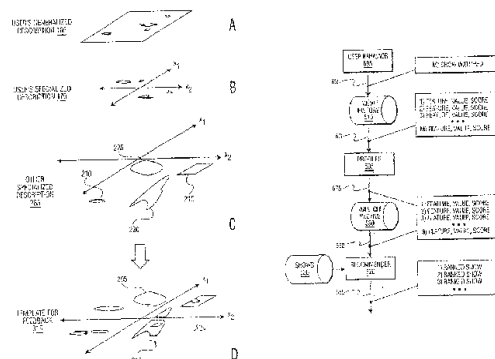
最終頁に続く

(54) 【発明の名称】 他人のプロフィールを用いた、個人用推薦装置のデータベース

(57) 【要約】

【課題】 ユーザの嗜好プロフィールにより提供された選択肢を、他人の嗜好、特に同じ家庭内のユーザの嗜好に基づいて拡張する機構を提供すること。

【解決手段】 テレビ番組を推薦する電子番組ガイドのようなデータ・クラスの推薦装置により、ユーザが同じ番組題材を何度も選択する場合、ユーザが型に陥ってしまうことがなくなる。一実施例の場合、別のユーザのプロフィールを利用してユーザのプロフィールを広げるように、推薦装置を自動的にプログラムすることができる。例えば、推薦装置は、ユーザと同じ家庭内の他のユーザの対象説明を、ユーザのプロフィールを広げるためのガイドとして用いることができる。これに代えて、家庭のプロフィールを、ユーザからのフィードバックを求めるためのソース題材用のフィルタとして用いることができる。こうして、単にユーザの範囲を恣意的に広げるのではなく、ユーザと何らかの形で関係のある他のプロフィールからの手本を得て、かつこれを利用する。この「関係」には、友人、ユーザの関心などを表す公けになっている定型概念などを含めることができる点に留意された



【特許請求の範囲】

【請求項 1】

様々なデータ・クラスに該当する例を評点する第一ユーザからフィードバックを受信するステップと、

当該第一ユーザのユーザ・プロフィールを当該フィードバックに応じてより正確にするステップと、

当該第一ユーザのユーザ・プロフィールを、第二ユーザのユーザ・プロフィールからのデータに応じて変更するステップと、

を有する、データ・クラスの推薦装置のために第一ユーザのユーザ・プロフィールを変更する方法であって、

当該変更ステップが、

少なくとも1つのデータ・クラスの推薦頻度が、他のいかなるデータ・クラスの推薦頻度も減少させることなく増加され、これにより、当該第一ユーザのユーザ・プロフィールの範囲が、当該第二ユーザのユーザ・プロフィール内に格納されている嗜好に従って拡張されるように変更すること、

を含む、方法。

【請求項 2】

当該第一ユーザのユーザ・プロフィールが、好ましいデータ・クラスの特定化された対象説明を含み、かつ、

当該変更ステップが、当該第二ユーザのユーザ・プロフィールの少なくとも1つの特定化された対象説明が含まれるように、当該特定化された対象説明を一般化すること、を含む

請求項 1 に記載の方法。

【請求項 3】

当該変更ステップが、少なくとも、当該第一ユーザのユーザ・プロフィールの特定化された説明と、当該第二ユーザのユーザ・プロフィールの特定化された説明との結合体を、当該第一ユーザのユーザ・プロフィールの当該特定化された説明の代わりにすることを含む、請求項 2 に記載の方法。

【請求項 4】

当該一般化ステップが、少なくとも、当該第一ユーザのユーザ・プロフィールの特定化された説明と、当該第二ユーザのユーザ・プロフィールの特定化された説明との結合体を、当該第一ユーザのユーザ・プロフィールの当該特定化された説明の代わりとすることを含む、請求項 2 に記載の方法。

【請求項 5】

当該第一ユーザのユーザ・プロフィールを訂正するためのテスト・データを、少なくとも1つの第二ユーザのユーザ・プロフィールからのデータに応じて選択するステップと、当該テスト・データに対する当該第一ユーザからのフィードバックを要求し、かつ当該第一ユーザのユーザ・プロフィールを当該フィードバックに応じて変更するステップと、を含む、請求項 1 に記載の方法。

【請求項 6】

当該選択ステップが、

テスト・データであって、これに対するフィードバックが当該第一ユーザのプロフィールに取り入れられて、当該第一ユーザのユーザ・プロフィールの区別力が向上するテスト・データ、

のみを選択することを含む、請求項 5 に記載の方法。

【請求項 7】

当該選択が、

テスト・データであって、当該テスト・データが好まれるかまたは好まれないかを当該推薦装置が決定するために、当該第一ユーザのユーザ・プロフィールでは不十分なテスト・データ、

10

20

30

40

50

を主に選択することを含む、請求項5に記載の方法。

【請求項8】

当該選択ステップが、概念領域の特定化された説明によってデータ選択範囲をフィルタすることを含む、請求項5に記載の方法。

【請求項9】

学習エンジンと、

当該学習エンジンに接続可能なユーザ・インターフェース装置と、

を有するデータ・クラスの推薦装置であって、

当該学習エンジンが、データ選択肢の説明を含んでいるデータ・ソースに接続可能であり、

当該学習エンジンが、当該ユーザ・インターフェースを介して、当該データ選択肢を評価する第一ユーザからのフィードバックを受信し、かつ、当該第一ユーザによって好まれているかまたは好まれていないデータ選択肢の説明を漸進的に生成することにより、第一ユーザ・プロフィールを生成するようにプログラムされ、

当該学習エンジンが、さらに、当該第一ユーザ・プロフィールに応じて当該第一ユーザのためのデータ選択肢の推薦を生成するようにプログラムされ、

当該学習エンジンが、さらに、当該第一ユーザ・プロフィールと、第二ユーザの少なくとも1つの第二ユーザ・プロフィールとに応じて、当該第一ユーザのためのデータ選択肢の推薦を選択的に生成するようにプログラムされる、

データ・クラスの推薦装置。

【請求項10】

当該第一ユーザ・プロフィールが、対象データ選択肢を定義する狭い説明と、対象外のデータ選択肢を定義する広い説明とを含み、

前記推薦が当該広い説明と当該狭い説明との間にある選択空間から導出されるように、当該学習エンジンがプログラムされる、

請求項9に記載の推薦装置。

【請求項11】

当該学習エンジンが、当該第一ユーザ・プロフィールが少なくとも対象のデータ選択肢を定義する狭い説明を含むようにプログラムされ、かつ、

当該学習エンジンが、さらに、当該狭い説明内の狭さのレベルを閾値と比較することにより、当該閾値より狭い目標データ範囲を含む推薦が当該第一ユーザ・プロフィールから得られるようにプログラムされ、かつ、

当該学習エンジンが、さらに、当該レベルを当該閾値に比較した結果に応じて、当該第一ユーザ・プロフィールと当該少なくとも1つの第二ユーザ・プロフィールとに応じて、当該第一ユーザのためのデータ選択肢の推薦を選択的に生成するように、プログラムされる、

請求項9に記載の推薦装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ユーザの行動を観察することによりユーザの嗜好を学習し、かつこの観察された嗜好に基づいて大きなデータ空間をフィルタする検索エンジンに関する。このようなシステムは、規則を明示的に入力することをユーザに要求するのではなく、ユーザの行動から規則を推定するアルゴリズムを用いる。本発明は、より具体的には、個人ユーザに対する推薦を、このユーザによる選択と他人による選択との両方に基づいて行う検索エンジンに関する。

【背景技術】

【0002】

検索エンジンは、非常に大きなデータベースを効率的かつ迅速に利用しなければならないアプリケーションにおいて、ますます重要になっている。検索エンジンは、ワールド・ワ

10

20

30

40

50

イド・ウェブの検索だけでなく、カタログ、テレビ番組、音楽一覧、ファイル・システムなどの格納にも効果的である。重点が情報から知識へと移行している世界において、検索エンジンは巨大な成長分野であり、かつ測り知れない可能性を秘めている。

【0003】

検索エンジンの用途が見い出される一分野として、ユーザの選択行動を観察し、かつこの行動に基づいて推薦を行う、いわゆる受動的な推薦装置が挙げられる。この技術は、テレビ番組を選択するための電子番組ガイド（EPG）と共に用いられる。

【0004】

電子番組ガイド（EPG）により、テレビと他のメディアとの無数にある視聴選択肢の中から選択を行う作業が、より管理し易くなることが見込まれる。受動的な検索エンジンは、ユーザ嗜好データベースを構築し、かつこの嗜好データを用いて提案を行って、現在または将来の番組情報をフィルタすることにより、選択を行うジョブを単純化したり、またはユーザに代わって選択を行うことさえする。例えば、システムは、ユーザからの特定の要求、またはユーザが推薦する選択肢が強調されていなくても、番組を記録することができる。

【0005】

上述したように、嗜好データベースを構築するためのデバイスの1つのタイプとして、ユーザの立場から見て受動的なタイプがある。ユーザは、単に、生のEPGデータから通常の方法で選択を行い、かつシステムは、ユーザの行動のモデルをこれらの選択肢から抽出することにより、個人的な嗜好データベースを徐々に構築する。次に、システムはこのモデルを用いて、このユーザが将来何を見ることを好むかを予測する。この抽出処理は、同じ項目に対して繰り返される要求を検出することによって、明らかに好きなものを識別するような、単純なアルゴリズムに従うことができる。または、この抽出処理を、多数の入力を用いた（自由度の大きい）決定木技法のような精巧な機械学習処理とすることができる。一般的な言い方をすると、このようなモデルは、ユーザの相互作用行動（すなわち、選択を行うためのユーザ・インターフェース（UI）による相互作用）の中にあるパターンを捜す。

【0006】

ユーザのテレビを見るパターンから有用な情報を抽出するための、容易かつ相当確実な技法の1つは、特徴-値（feature-value）の計数表を生成することである。特徴の一例として「一日の時刻」があり、かつこれに対応する値を「朝」とすることができる。ある選択が行われると、この選択を特徴付けている特徴-値の計数が増加する。ある所定の選択肢は、通常、多くの特徴-値を有するであろう。この選択肢がそこから区別された番組のサブセットを（オプションで同時に）選択することにより、1組の否定的な選択肢を生成することができる。これらの番組の各々の特徴-値の計数は、減少するであろう（すなわち、見られなかった番組の計数は増加するであろう）。これらのデータは、ベイジアン予測装置に送信される。このベイジアン予測装置は、これらの計数を、候補を特徴付けている特徴-計数への重みとして用いることにより、ある候補がユーザの嗜好に合う確率を予測する、このタイプのプロファイリング機構は、2000年2月4日に出願された米国特許出願第09/498271号「ベイジアン・テレビ番組推薦装置（BAYESIAN TV SHOW RECOMMENDER）」に解説されている。この特許出願の全体は、本明細書に完全に解説されているかのように、本明細書に参照として取り入れられている。さらに、ユーザ行動の観察例からプロフィールを受動的に構築する、これと同じシステム・クラスの規則に基づく推薦装置が、1999年1月14日に公開されたPCT出願、国際公開番号第W099/01984号「インテリジェント電子番組ガイド（INTELLIGENT ELECTRONIC PROGRAM GUIDE）」に解説されている。

【0007】

第一タイプの別の例として、MbTVがある。これは、視聴者の視聴パターンをモニタすることにより、視聴者のテレビを見る嗜好を学習するシステムである。MbTVは、透過的に動作し、かつ視聴者の好みのプロフィールを構築する。このプロフィールを用いて、例えば、視聴者が見たいと関心を抱く可能性があるテレビ番組を推薦するサービスが提供される。

MbTVは、その視聴者の各好みを学習し、かつ学習したことを利用して、放送が近づいた番組を推薦する。MbTVは、放送が近づいた所望の番組に対して視聴者の注意を促すことにより、視聴者がテレビを見る予定を組むことを助けることが可能となる。また、格納デバイスを加えることにより、視聴者の不在時に、これらの番組を自動的に記録することができる。

【0008】

MbTVは、嗜好決定エンジン(Preference Determination Engine)と格納管理エンジン(Storage Management Engine)を有する。これらは、テレビの留守録画を容易にするために用いられる。MbTVは、所望の番組を単に提案するのではなく、自動的に記録することができる。MbTVの格納管理エンジンは、格納装置が最適のコンテンツを有している状態を確保することを試みる。このプロセスには、記録済みの番組のどれが(完全にまたは部分的に)視聴されているか、かつどれが無視されているかを追従することが含まれる。視聴者は、今後視聴するための記録済みの番組を「ロック」して、消去を防止することができる。提案された番組または記録済みのコンテンツを視聴者がどのように扱ったかが、この情報を用いて将来の決定をより正確化するMbTVの嗜好エンジンに、さらにフィードバックされる。

10

【0009】

MbTVは、各「関心要素」を表すための、記録空間の一部を予約するであろう。これらの「関心事」は、異なる家族員に移しても良いし、または異なる好みの部類を表すこともできる。MbTVは、ユーザの介入を必要としないが、その機能を微調整することを望むユーザによってカスタム化することができる。視聴者は、異なるタイプの番組に対する「格納配分(storage budget)」に影響を及ぼすことができる。例えば、家庭内でテレビの大部分を見ているのが子供であるにも関わらず、子供向け番組により消費される記録空間を25%以下にする必要があることを、視聴者は指示することができる。

20

【0010】

第二タイプのデバイスは、より能動的である。このデバイスにより、ユーザは、特徴を類別化することによって、好きな番組または嫌いな番組を指定することができる。これらを、特徴-値の対の評点化(この特徴に対する重みに、ある値を加算したもの;例えば、重み = 特徴の重要度 + 嗜好度の値または不評度の値)、または、お気に入りの番組や、「ドキュメンタリーは好きだが、ギャングものが放送される木曜日の夜は除外する」のような特徴-値の対を組み合わせたような何か他の規則を指定したもの、とすることができる。例えば、ユーザは、ドラマとアクション映画が好きで、かつ特定の俳優が嫌いであることを、ユーザ・インターフェースを介して示すことができる。したがって、これらの基準を適用することにより、1組の番組の中からどれがユーザによって好まれるかを、予測することができる。

30

【0011】

第二タイプのシステムの一例として、欧州特許出願第0854645号は、ホーム・コメディ、ドラマ・シリーズ、古い映画などの好ましい番組部類のような包括的な嗜好を、ユーザが入力することができるシステムを解説している。この出願は、例えば、あるプロフィールは10~12歳の子供用、別のプロフィールは十代の女の子用、別のプロフィールは飛行機愛好者用などとした嗜好プロフィールを選択することができる嗜好テンプレートも解説している。

40

【0012】

第三タイプのシステムは、ユーザが、ある方法で番組をランク付けすることを可能にする。例えば、現在、TIVO®は、ユーザが番組に対して、良いという評価を3つまで、または悪いという評価を3つまで与えることを可能にしている。この情報は、達成可能な特徴-値の対に与えられる加重分解能の程度をより細かくすることができる点を除き、第二タイプのシステムに幾つかの点で類似している。また、ユーザの好みをこの情報によって表現した方がより明確な点を除き、第一タイプに類似している。(ただし、このことは、第三タイプのシステムのような、ユーザによるランク付けと組み合わせされた米国特許出願第09/498271号に解説されているペイジアン技術が従来技術であることを、認めるものではな

50

いことに留意されたい。))

【0013】

PCT出願「テレビの予定情報を用いるためのシステムと方法 (System and Method for Using Television Schedule Information)」(国際特許公開公報第W0 97/4924号)は、第三タイプの一例である。これは、ユーザが、通常の格子様式で表示された電子番組ガイドにより操作を行い、かつ様々な番組を選択することができるシステムを解説している。各点では、記録または視聴するための番組を選択すること、番組を見るための備忘録を予定組みすること、かつ番組を選択してお気に入り指定することを含む、解説されている様々なタスクをどれでも行って良い。ある番組をお気に入り指定する目的は、おそらく、「この番組を見るオプションを常に表示」したり、または備忘録を繰り返し実施するような、確定している規則を実施することである。お気に入り指定する目的は、この出願には明確に説明されていない。しかしながら、より重要な点は、嗜好データベースを作成する目的のため、ユーザが番組を選択してお気に入り指定する際に、この番組がお気に入りである理由を示すオプションをユーザに提供することができることである。この理由は、包括的な嗜好を定義することにより、他の明示的な基準と同様に示される。

10

【0014】

第一タイプのシステムの利点は、ユーザが明示的なデータを提供する必要が全くないので、ユーザにとってより容易なことである。ユーザは、システムと相互作用するだけで良い。様々な機械学習方法または予測方法のどれも効果的にするには、相互作用のかなりの履歴を利用可能にして、有効な嗜好データベースを構築しなければならない。第二タイプと第三タイプの利点は、明示的な嗜好情報を提供することである。第二タイプには信頼性があるが、ユーザは、どの基準がディスクリミネータとして良いか、かつこれらにどんな重み付けをすべきかが決定可能となる程には、自分自身の嗜好を抽象化することが出来ない場合があるので、完璧ではない。第三タイプは、ユーザの負担とならず、かつおそらく、情報の質は最良だが、第二タイプにより得ることができるすべての情報を生成するにも関わらず、これらの情報をすべては含めることが出来ず、かつさらに、第一タイプのように多くの番組に関する情報を必要とする場合があることは、負担となり得る。

20

【0015】

嗜好データベースを構築するための従来技術の問題の1つは、ユーザが同じ番組を繰り返し見る場合に発生する。ユーザの選択肢のかなりの割合は、これらの選択肢から抽出された、小さ過ぎるデータの組により構成されるので、これらの規則から抽出された規則は、結局、範囲が過度に狭い推薦を定義することになる。この問題は、ある型に陥ることに似ている。従来技術の別の問題は、暗黙のプロフィールをユーザ間で容易に共有することが出来ないことである。ユーザが、友人が推薦した番組を好きな場合でも、このユーザが友人のプロフィールのある部分またはすべての部分を得て、かつこれを、ある方法で自分のプロフィールに組み合わせるための良い方法がない。

30

【0016】

【特許文献1】

米国特許出願第09/498271号

【特許文献2】

国際特許公開公報第W0 97/4924号

40

【発明の開示】

【課題を解決するための手段】

【0017】

本発明は、ユーザの嗜好プロフィールにより提供された選択肢を、他人の嗜好、特に、同じ家庭内のユーザの嗜好に基づいて拡張する機構を提供する。肯定的および/または否定的な例に基づいて、選択エンジンを生成し、かつさらに正確化するための様々なタイプの機構が知られている。バージョン空間アルゴリズムと呼ばれているタイプの機構は、データベース(すなわち「選択空間」)内の利用可能なすべての可能な選択肢に関する2種類の説明を保存する。

50

(1) 選択空間の最も広い説明である一般的な説明であって、否定的な選択肢がすべて除外されている。

(2) 選択空間の最も狭い説明である特定化された説明であって、選択空間内の肯定的な例をすべて含んでいる。

否定的または肯定的な例が提供される度に、この例を用いて、特定化された説明または一般化された説明が相応に変更される。このアルゴリズムと、バージョン空間アルゴリズムに関するさらなる詳細は、米国特許出願番号第09/794445号「番組内容の一般化と専門化によるテレビ番組の推薦」に記述されている。この米国特許出願は、本明細書にすべて完全に説明されているかのように、本明細書に参照として取り入れられている、

【0018】

一般化された説明は、テレビ番組の選択範囲内で、ユーザが関心を抱く可能性のあるすべての可能な番組選択を示す。特定化された説明は、ユーザが明らかに興味を抱くすべての可能な番組説明を示す。一般化された説明と特定化された説明との間の説明範囲は、大きなものとなり得る。さらに、一般化された説明は自由度が高くなり過ぎて、選択肢の大きな組を妥当な数に縮小することが出来なくなる可能性があり、かつ特定化された説明は狭過ぎて、狭い範囲の例が捕らえることが出来なくなる可能性がある。

【0019】

従来技術は、この混乱状態からユーザを解放するための他の方法を提供している。1つの方法は、一般化された説明により定義された大きな空間から番組コンテンツを無作為に選択し、かつこれらをランク付けすることをユーザに求めることである。しかしながら、これは、かなり無分別な営みとなり得る。例えば、提供されている唯一の例が英語による例であると仮定する。ユーザは、英語以外の説明空間内にあるコンテンツの否定的な例を挙げていない。しかしながら、大半のユーザは、テレビを見ることによって言語範囲を広げることには、気が進まない可能性がある。したがって、無作為の選択装置は、英語空間外で例を捕らえ、かつこれらの例をランク付けすることをユーザに求めるであろうが、僅かに有効な基準しか得られない。つまり、この例が車に関するもの、またはスペイン語だったために、ユーザは、この例が気に入らなかったのであろうか。ユーザは、余りに多くの無関係な選択肢をランク付けすることを求められた場合、直ぐに退屈してしまうであろう。ユーザの一般化された説明よりも、より狭い説明から例を引いた方がよいだろう。本発明によると、これは、特定化された説明、またはこのユーザとある基準で類似した他人（例えば、同じ家庭内のユーザ）の説明に影響を与えることにより、行うことができる。

【0020】

一実施例の場合、ユーザにより選択された1人以上の他人の特定化された説明の全空間を含んだ一般化された特定化された説明が定義される。この一般化された特定化された説明は、ユーザの肯定的かつ否定的なフィードバックがそれに対して求められるテスト・サンプルを生成するためのソース・フィルタとして用いられる。他の一実施例の場合、家庭内の全ユーザのような一群が自動的に定義され、かつすべての特定化された説明により定義される空間を含んだ、最も狭い新たな特定化された説明が生成される。テスト・サンプルは、同様にこの新たな特定化された空間から導出される。

【0021】

上記の両実施例の改良例の場合、ユーザの特定化された説明内の曖昧な特質を区別化するテスト・サンプルが優先される。すなわち、ユーザの特定化された説明に適合する、一般化された特定化された説明からのサンプルは既に回避されており、かつこの説明以外のサンプルは好まれる。ユーザの特定化された説明が、それに沿って、一般化された特定化された説明と融合する次元では、後者のサンプルの方が、明らかにより高い区別力を有する。

【0022】

上記の方法の別の改良例は、ユーザの一般化された説明を用いて、一般化された特定化された説明を特定化することである。この一般化された説明は、ユーザが好きではないものの貯蔵庫なので、一般化された特定化された説明の空間をフィルタするためのフィルタと

10

20

30

40

50

して用いることができる。

【0023】

別の実施例の場合、ユーザの各クラスが定義され、かつ、共同フィルタリングに似た方法で、ユーザの特定化された説明が一般化され、原型となるユーザの特定化された説明の空間が含まれる。例えば、サービス・プロバイダは、「スポーツ狂」、「血液と内臓」、「歴史おたく」、「感傷的で感情的なもの」、「科学愛好家」および「ファンタジー愛好家」のような定型概念に対する特定化された説明を生成することができる。

【0024】

別の実施例の場合、ユーザの説明をより正確化するフィードバック用のソースを作成するために、他の特定化された説明を用いるのではなく、他の特定化された説明に影響を与え、ユーザの特定化された説明の代わりとなる。

10

【0025】

ユーザの特定化された説明が、一般化された特定化された説明の代わりとなる実施例を支援するユーザ・インターフェースの場合、定型概念の試用を、ある期間、ユーザに求めることができる。ユーザが結果を気に入らない場合、古い特定化された説明を回復することができる。オプションで、ユーザは、ユーザの特定化された説明を一般化するために定型概念的な説明が適用されている間に得られるフィードバックによる利点を、維持することができる。

【0026】

本発明は、他のタイプの誘導エンジンに拡張可能である。例えば、他のネットワークからの予測に基づいてニューラル・ネットワークをトレーニングし、好き嫌いに関するニューラル・ネットワークの予測を一般化することができる。決定木は、別の決定木により生成されたサンプルを加えたり、またはより直接的に別の決定木から枝を共有するような公知技術により拡張可能である。他のタイプの機械学習も、まだ未知のものでさえ、本発明の背後にある基本的な概念を利用することができるので、本出願の教示内容と組み合わせれば、当業者の能力の範囲内となるはずである。

20

【発明を実施するための最良の形態】

【0027】

本発明をより完全に理解することができるように、例示的な図を参照しながら、本発明を特定の好ましい実施例に関連付けて説明する。これらの図に関して強調すべき点は、図示されている特定事項が、例示的なものであり、かつ本発明の好ましい実施例を例示的に解説することしか目的としておらず、かつ本発明の原理と概念の側面に関する最も有効的で、かつ容易に理解される説明と考えられるものを提供するために示されていることである。この点に関して、説明を図面と共に理解することによって、本発明の幾つかの形態を実際に実施可能な方法が当業者に明らかとなるので、本発明の構造的な詳細を、本発明の基礎的理解に必要とされるよりも詳しく示す試みはされていない。

30

【0028】

図1を参照すると、概念空間100が、説明上の形式表現に従って定められている。例えば、図1は、各フレーム・スロット内の値に対してベン・タイプの表現を用いた、フレームに基づいたデータ構造または表現言語を示している。説明のため、フレームに基づいた構造内の多数のスロットは、フレームに基づいた構造内のスロットのような記述子構成要素を表す、2本の軸(x_1 と x_2)として表示されている。選択されたスロットはどのようなパラメータを表しても良く、かつこの線図の目的は、パラメータが独立しているか、またはパラメータの数が制限されていることを示すことではないことを、理解されたい。例えば、軸 x_1 は、テレビ番組のタイプ(コメディ、ドラマ、ホラー、スポーツなど)を表示し、かつ x_2 は、俳優(トム・クルーズ、シェリー・デュバル、ロバート・ワグナーなど)を表すことができる。説明のため、多くの異なる記述子構成要素があって、これらの各々が、値または値の範囲を1つ以上有していても良く、かつ別の記述構成要素に依存していても依存していなくても良い、ということを想像することができる。

40

50

【 0 0 2 9 】

可能な説明領域（概念空間100）は、この形式表現に固有の偏りのみにより制限される。ここで、あらゆる可能な説明が、概念空間の最も高いレベルの空白の一般化された説明115に含まれている。何かの学習が行われる前に、この単集合の一般化された説明115に、あらゆる可能な例が含まれる。概念空間の最も低いレベルには、ユーザにより提供された最初の肯定的な例しか含まない単集合130がある。

【 0 0 3 0 】

肯定的な例と否定的な例とを用いて、ある期間トレーニングが行われた後、例えば、参照として取り入れられている上記の出願に解説されているバージョン空間アルゴリズムを利用して、もっとも最近の特定化された説明170が広げられ、この結果、この説明170は、肯定的な例をすべて含んだ最も狭い説明の組となる。もっとも最近の特定化された説明170は、本質的に、否定的な例をすべて除外する。さらに、トレーニングの後、否定的な例をまったく含んでいない最も広い可能な説明の組である一般化された空白の説明115から、最新の一般化された説明165が導出済みとなる。この最新の一般化された説明165は、本質的に、肯定的な例をすべて含んでいる。

【 0 0 3 1 】

最新の特定化された説明170により定義されている選択空間からの選択肢には、以前の肯定的な例と類似した選択肢しか含まれていない。したがって、最新の特定化された説明170から推薦が導出される場合、これらの推薦は狭くなり過ぎてしまうので、ユーザは、狭過ぎる1組の例に肯定的なフィードバックを与えてしまうことにより、自分の型にはまり込んでしまうであろう。このような場合、ユーザは、広過ぎる一般化された説明を有することもあり得るので、この一般化された説明が、広過ぎる空間に拡張されてしまう場合がある。これらの極端な場合の間には、バージョン空間101と呼ばれる空間がある。これは、ユーザが好む可能性がある対象事項に対する可能な説明を定義したもので、一般化された説明から特定化された説明に移動するにつれて、確実性が増す。

【 0 0 3 2 】

次に、図2A～2Cを参照すると、ユーザの特定化された説明280と別の特定化された説明285とを合わせることで、新たな特定化された説明290が導出されている。別の特定化された説明285は、例えば、定型概念的な説明または別のユーザの説明で良い。ここで、範囲110、115、120、125の統合体であるユーザの組に、範囲210、215、220、225の統合体である他の組が組み合わされる。この結果、図2Cに示されている、隣接する範囲250、255、260、265、270、215の統合体により定められた組が得られる。より正確には、この新たな説明とは、ユーザの特定化された説明280が、他の特定化された説明285に含まれている対象事項を除外しないように生成されたものである。この一般化された特定化された領域には、このユーザと同じ家庭内の他のユーザの、複数の他の特定化された領域が含まれることが好ましい点に留意されたい。家庭内の他のユーザと一貫した方法で拡張することにより、ユーザ自身のプロフィールよりも良好な予測が可能となることが見い出されている。

【 0 0 3 3 】

型に陥ってしまっているプロフィールを拡張するために追加的なユーザ・プロフィールを用いることを、ユーザが選択できるようにすることができる。ユーザのオプションを拡張するために用いられる、一群のユーザ・プロフィール、定型概念化されたプロフィール、または一つ以上の特定のプロフィールを選択するオプションを、ユーザに提供することができる。他のプロフィールを用いて、ユーザのプロフィールを永久に変更したり、または単にユーザ毎に選択範囲を広げて良い。また、学習エンジンが、ユーザのプロフィールが型に陥ってしまったことを検出し、かつ、家族員すべての特定化された説明を加えることにより修正処置を行うことも可能である。これは、プロフィールのタイプに応じて、様々に決定することができる。例えば、特徴-値-得点-タイプのプロフィールの場合、特徴-値-得点の記録を僅かしか有さないプロフィールを、型にはまっているとして識別することができる。概念空間内では、非常に特定化されている特定化された説明なら、このプロフィールが型にはまっていることを示すだろう。同じ年齢の家族員を区別し、かつこれら

10

20

30

40

50

の家族員が同様の年齢部類にいる場合しか、説明を共有しないことが適切な場合があることに、留意されたい。

【0034】

従来技術で公知のように、システムは、無作為に選択された新たな例に対するフィードバックを求めることができる。しかしながら、このような方策は、否定的なフィードバックが返された題材が含まれてしまう可能性があり、かつ大き過ぎる可能な対象事項の空間が含まれてしまう可能性がある。大半が否定的な例が見つかり、ユーザが失望し、かつ関心をなくしてしまう可能性が高い。これに代えて、最新の一般化された説明165を、新たな例に対するフィルタとして用いることができる。しかしながら、最新の一般化済みの説明165は、実用化するには大き過ぎ可能性のある空間を、やはり定義してしまうかも知れない。

10

【0035】

この問題に対する方法の1つは、別のユーザの特定化された説明をフィルタとして用いて、フィードバックを求めることである。システムは、別のユーザのプロフィールの特定化された説明を、新たな題材を選択するためのフィルタとして用い、かつこの新たな題材に対するユーザのフィードバックを求めても良い。図3A~3Dを参照すると、ユーザが既にフィードバックを返している題材をテスト例から除外することが好ましい。このように、ユーザの一般化された説明165内とユーザの特定化された説明170内の対応する部分を、他の特定化された説明285から除去して、フィードバック用の新たなテンプレート315を設けても良い。これらの図中には、他の特定化された説明170が1つしか示されていないが、任意数の特定化された説明の統合体を用いて、フィードバック用のテンプレートが生成可能であることは明らかである。

20

【0036】

ユーザが他人のプロフィールを利用して、ユーザ自身のプロフィールを拡大することが可能になることに関する重要な論点の1つは、このプロセスに対する支配感覚をユーザに与えてしまうことである。ここで、主な懸念事項は、おそらく、他人が行う可能性があることが、ユーザに明らかになってしまうことだろう。他人のプロフィールに影響を与えることを、透過的に行って良い場合もある。例えば、推薦装置は、ユーザの個人プロフィールにしか依存しないのではなく、このユーザと同じ家庭内の他のユーザのプロフィールから導出された推薦を含めても良い。これを、ある時間だけ、または常時行うことができる。勿論、フィードバックが得られた場合は、常にこのフィードバックを用いて、個人ユーザのプロフィールをより正確化しても良い。

30

【0037】

上記の説明は、バージョン空間アルゴリズムにより示されている比喩的な用語と図を用いているが、本発明は、他のタイプの推薦装置システムにも適用可能である。第一ユーザが、別のユーザのプロフィールにより推薦された例を気に入ったと仮定する。第一ユーザが他のユーザのプロフィールを用いて、自分のプロフィールを変更することを可能にする方法の1つは、他のユーザのプロフィールを用いて生成され提案された番組に対して、他のユーザのプロフィールを用い、かつ第一ユーザがこれらの番組にフィードバックを提供できるようにすることである。これは、推薦エンジンとの互換性がこれらの番組にまったくなくても、行うことができる。

40

【0038】

ユーザのプロフィールを拡張するための別の方策は、別のユーザの一般化された説明を、ユーザの一般化された説明の代わりにすることである。

【0039】

図5を参照すると、映像信号470を受信し、かつチャンネル変更機能を制御し、かつユーザが、テレビのチューナ430ではなく、コンピュータ440にリンクされたチューナ445によるチャンネル選択を可能にするために設けられたコンピュータ440が、本発明を支援することができるハードウェア環境の例に含まれている。ユーザは、次に、コンピュータを制御する遠隔制御装置410を用いて、表示された番組予定から所望の選択肢を強調することにより

50

、視聴すべき番組を選択することができる。コンピュータ440は、更新済みの番組予定データを、コンピュータ440がそれを介して受信することができるデータ・リンク460を有する。これを、インターネット・サービス・プロバイダに接続可能な電話線、または他の適切な何らかのデータ接続とすることができる。コンピュータ440は、番組予定情報、番組用途と更新内容、および他の情報を格納するための、大量格納装置435（例えば、ハードディスク）を有する。ユーザの嗜好に関する情報と他のデータは、メモリ・カードまたはディスク420のような取り外し可能な媒体を介して、コンピュータ440にアップロード可能である。

【0040】

上記の例示的なハードウェア環境の場合、多くの代替例が可能であり、かつこれらはすべて、本発明に関連付けて使用可能であることに留意されたい。大量格納装置は、揮発性メモリまたは不揮発性メモリに置換可能である。データは、局所的または遠隔的に格納可能である。事実、コンピュータ440の全体を、リンクを介して離れて動作するサーバに置換することができる。遠隔制御装置は、赤外線ポート415を介したコンピュータ440への指令送信を行うために使用するのではなく、映像を伝える物理チャネルから分離させたり、または映像を伝える物理チャネルと同じにすることができるデータ・チャネル460を介して、指令を送信することができる。映像470または他のコンテンツは、ケーブル（RF）、または他の任意のブロードバンド物理チャネルで伝えたり、または大容量記憶装置または取り外し可能な格納媒体から得ることができる。映像470または他のコンテンツは、電話回線のようなスイッチされている物理チャネル、または同期データ通信に適したATMまたは他のネットワークのような仮想的にスイッチされているチャネルにより、伝えることができる。コンテンツを非同期化し、かつドロップアウトに対する耐性を有することにより、現代のIPネットワークが利用可能となる。さらに、番組のコンテンツがそれを介して受信されるラインのコンテンツは、音声、チャット対話データ、ウェブサイト、または様々な選択肢が可能な他のいかなる種類のコンテンツとすることができる。番組ガイド・データは、独立したデータ・リンク460以外のチャネルを介して受信可能である。例えば、番組ガイド情報は、映像または他のコンテンツと同じ物理チャネルを介して受信可能である。メモリ・カードまたはディスク420のような取り外し可能なデータ格納媒体を介して、番組ガイド情報を提供することもできる。遠隔制御装置410は、キーボード、音声指令インターフェース、3Dマウス、ジョイスティック、または他の任意の適切な入力装置に置換可能である。強調用インジケータを移動したり、選択肢を記号（例えば、名前または数）で識別したり、または、データ伝送または取り外し可能な媒体を介したバッチ形態での選択を行うことにより、選択を行うことができる。取り外し可能な媒体の場合、1つ以上の選択肢をある形態で格納し、表示装置170を完全にバイパスさせてコンピュータ440に伝送することができる。例えば、バッチ・データは、携帯型格納デバイス（例えば、携帯情報端末、メモリ・カード、またはスマート・カード）からロード可能である。このようなデバイスは、様々な環境で用いるための多くの嗜好を格納して、使用されるコンピュータ装置をカスタム化することができる。

【0041】

幾つかのタイプのプロファイリング機構は、これらの内部にある対象説明を、抽象的説明として表示することができる。例えば、フレームに基づいたデータ構造の場合、異なるスロットにタイトルを関連付けることにより、1ユーザが別のユーザのプロフィールを調べることが実際に可能となる。スロットは独立していないので、任意の1つのスロットを選択したことによる影響が、他のスロット内で可能となっている選択肢に影響を与える可能性があるが、プロフィールの構成方法の有意な見方をユーザに示すことは、必ずしも容易な作業ではない。例えば、俳優のトム・クルーズがユーザのお気に入りであることを示す特定化された説明を、ユーザのプロフィールに含めても良い。しかしながら、肯定的なフィードバックが与えられた例は、アクション・タイプの映画に限定されている。したがって、ユーザがトム・クルーズが好きだと言うことは出来ない。ユーザは、特定タイプの映画でのトム・クルーズしか好きではない、ということかも知れない。上記の例は、単純な

10

20

30

40

50

ものである。実際の実施例は、非常に複雑となる可能性があるので、ユーザに示すことは困難となる。インターフェースは、複数パラメータ空間を定義している関心事のスロットを有するリンクされているスロットを、すべて示さなければならないだろう。しかしながら、正確度を100%にすることが目標ではないと考えるべきである。目標は、単に、ユーザが他のユーザのプロフィールのある側面しか借りることが出来ないようにすることで良く、かつこの側面の特徴付けをそれ程完全なものにする必要はない。システムは、たった1本のスロット内の値に基づいて変更タグを付けることにより、多くの他のスロットに結合された特定のスロットに基づいて、ユーザのプロフィールの変更を行うことができる。したがって、第二ユーザのプロフィールがトム・クルーズに対する著しい嗜好を示していることをシステムが第一ユーザに示した場合、第一ユーザは、この嗜好に基づいて自分のプロフィールを変更することを承諾する際に、自分のプロフィールを拡張することができる。この結果、システムは、第二ユーザのプロフィールに内在する付随的な予告記載すべてに結合されているトム・クルーズの例を推薦することが可能となる。換言すると、示されている例の場合、第一ユーザに対して、トム・クルーズを望んでおり、かつトム・クルーズを見たいがアクション映画でのトム・クルーズに限定するか、という問い合わせが行われる。

10

【0042】

バージョン空間アルゴリズムの下で条件付けられた、フレームに基づいたデータ構造の場合、ユーザのプロフィールの特徴を決定付ける「トム・クルーズ」のようなラベルは、ある値（例えば、他のスロット内の値と何度も組み合わせられているように見える「トム・クルーズ」）を選択することにより識別可能となる。換言すれば、このスロット値が特定化された説明内で発生する率は高くなる。図4Aと図4Bには、1つのプロフィールから別のプロフィールへの説明情報の移植をユーザが制御することを可能にするためのこの機構が示されている。ここで、例えば、ユーザの特定化された説明とすることができるユーザの説明がスキャンされ、かつその様々な部分が、主な特徴に従ってラベル付けされる。図示されているのは、部分210への「トム・クルーズ」というラベル付けである。比喩的に説明すると、データ構造 x_1 の一方の次元は、俳優に対応させて良い。他方の次元 x_2 は、映画または何か他のもののタイプのような他のパラメータに対応すると見なして良い。「トム・クルーズ」という値は、他のパラメータの複数の値に関連付けられて選択されているので、重要な特徴-値であると推定することができる。

20

30

【0043】

説明部分210は、他の図中の他部分と同様、隣接した閉空間として示され、近接した範囲を示しているが、このような特徴は、データが対象説明内でどのように表されるのかを表していてもいなくても良い点に留意されたい。フレームに基づいたモデルの場合、各特徴または各スロットは、別々の値を有していても良く、かつデータ・セットが210のような閉空間を形成する傾向を示すように、近接した特徴間には何の関係もなく良い。これは、説明のために取り入れた、単なる抽象概念である。この閉空間の唯一の側面は、「トム・クルーズ」という値が、その重要性を示す次元 x_2 に沿った他の特徴の複数の値に関連付けられている事実が、閉空間の330と示されている寸法の長さにより示されていることである。

40

【0044】

他のタイプのデータ構造の場合、プロフィールの各部分にラベルを付ける機構は、容易に識別されるだろう。例えば、特徴-値の対を格納するシステムの場合、重要な特徴にラベルを付け、かつこの特徴を別のプロフィールに移植することが、さらに容易となる。図6を参照すると、このようなシステムの場合、ユーザは、ある選択を好きまたは嫌いとランク付けするためにフィードバックを提供し、かつオプションで、好きまたは嫌いの程度を含める。例えば、システムは1~7の得点を用い（4を、どちらでもない、とする）、1~3で嫌いの程度を表し、かつ5~7で好みの程度を表すことができる。ユーザ・インターフェース（UI）500は、番組を一覧にし、かつフィードバック情報を受け入れるために用いられる。これに代えて、UI 500を、番組が終了した場合またはユーザが他の番組に切り換え

50

た場合に、番組に対するフィードバックを返すことをユーザに求める単純なプロンプトとしても良い。このプロンプトのタイプは、ある状況またはあらゆる状況において、必要な場合、ユーザがプロンプティングを無効化することを可能にする嗜好の組に従属することが好ましい。

【0045】

フィードバックUI 500の各例により生成された情報は、その選択に関連付けられた得点を有する、1つ以上の選択肢555（テレビ用データベースの場合は、番組）となる。この情報を用いて、このようなエントリを多数含むことができるフィードバック履歴ファイル505が満たされる。次に、フィードバック・データ560を、プロファイル550に適用することができる。これに代えて、このデータを最初にプロファイル550内で圧縮し、かつ次にフィードバック・プロフィール・データベース525内に格納することにより、圧縮形式で格納することができる。この圧縮データを、2000年2月4日に出願された米国特許出願第09/498 271号「ベイジアン・テレビ番組推薦装置（BAYESIAN TV SHOW RECOMMENDER）」に解説されているランク付けを各々が有する、1組の特徴-値の対565としても良い。選択が行われると、対応する得点を有する幾つかの（M個の）特徴-値の対565を生じることができる。肯定的なフィードバックと否定的なフィードバックとの両方が得られるように、ユーザは、好きな番組と嫌いな番組の両方を評価することが好ましい。例えば、見るために選択された番組にしかフィードバックが提供されないために、肯定的なフィードバックしか得られない場合、否定的な要因はデータベースに存在することが出来ない。このことは、この選択が行われたのと同時に、利用可能な番組のサブセットを選択することにより、1組の否定的な選択肢をシステムに生成させることによって、改良することができる。述べられているように、ユーザが、肯定的なフィードバックと否定的なフィードバックとを均等に提供することにより、否定的な選択肢を自動的にサンプリングする必要をなくすことが好ましい。これらの各々の特徴-値の計数は、減少するだろう。多くの選択肢に渡って格納されたこのデータを、フィードバック・プロフィール・データベース525内に格納することができる。次に、番組データベース520から導出された候補一覧に基づいて推薦装置580が推薦を行うと、N個の記録全体555が利用可能となる。このプロセスの結果、最終的に、番組データベース520から得ることができる選択肢がフィルタまたは分類された一覧575が得られる。この推薦装置は、ベイジアン・フィルタまたは他の任意の予測装置で良い。

【0046】

図7を参照すると、図6と非常に類似したプロセスを用いて、特徴-値の対によるプロフィール・データベースを生成することができる。この予測装置は、背景技術の部分で説明されている第一タイプのものである。ここで、ユーザが選択する番組選択肢が推定され、番組選択肢に対する肯定的な得点が示される。ユーザが選択を行うと、結果的に、オプションで得点が添付する、特定の番組665が得られる。この結果には、ユーザの反応の仕方から推定される得点を含めることもできる。ユーザが番組を最後まで見た場合、得点を上げ、かつ短時間しか見なかった場合、得点を下げて良い。番組がこれらの二つの間の期間だけ見られた場合、得点の大きさを中程度にすることができる。これに代えて、見られた番組は正の得点を受け、かつ見られなかった番組を無作為にサンプリングしたものは、（オプションで同時に）負の得点を得るようにしても良い。

【0047】

視聴履歴データベース510は、番組と得点を格納する。記録670は、得点が添付された特徴-値の対675を生成する、プロファイル595に供給される。得点が添付された特徴-値の対675は、内在的プロフィールのデータベース530内に格納することができる。内在的プロフィールのデータベース530のコンテンツ680は、次に、推薦装置620による利用が可能となる。推薦装置620は、これらのコンテンツを、最新の番組520からのデータに組み合わせ、推薦685を生成する。

【0048】

このタイプのプロファイルの場合、特徴を結合しないことにより、1つのプロフィールから別のプロフィールに移植可能なデータ部分へのラベル付けの問題は単純化される。した

10

20

30

40

50

がって、「俳優」という特徴と「トム・クルーズ」という値は、対象のプロフィール内で目立つようになるので、識別が容易になる。なぜならば、その特徴-値の対には、高得点が関連付けられるからである。ユーザには、別のユーザのプロフィールのその側面を選択して自分のプロフィールに移植するオプションを提供することが可能となる。この結果、ユーザのプロフィール内の対応する特徴-値の対に関連付けられている得点は、調整されるであろう。

【0049】

プロフィールが型にはまっているユーザを広げるために、特徴-値-得点タイプのデータを組み合わせることは、型にはまっていないユーザのデータベース内では高得点を有する特徴-値の対の得点を、型にはまっているユーザのプロフィール内で高める問題となるだろう。この場合も、型にはまっているユーザが、変更すべき特徴-値を選択することができるように、ユーザ・インターフェースを生成することができる。これに代えて、ユーザは、これを目に見えないように行わせることができる。別の代替手段は、この変更を一時的にしか行わないことにより、この変更を試みるということが可能である。型に陥ってしまう問題に対処する別の方法は、ユーザのプロフィールに関連付けられている非常に高い得点を何れも調整することである。ユーザは、これを選択的に行うことができる。ユーザ・インターフェースはユーザに対して、どの特徴-値が、（肯定的かまたは否定的か何れかの）非常に高い得点を有し、かつユーザによる変更が可能かを示すことができる。

10

【0050】

本発明は上述の例示的な実施例の詳細に限定されず、かつ本発明を、本発明の趣旨または本質的属性の範囲内で他の特定の形態で実施しても良いことが、当業者に明らかとなるだろう。したがって、本実施例は、すべての点において例示的かつ非限定的であると見なされ、本発明の範囲は、上述の説明ではなく、添付の請求の範囲により示されており、かつしたがって、これらの請求項と等しい意味と範囲内にあるすべての変更は、これらの請求項に含められるものとする。

20

【0051】

例えば、本発明を、テレビ推薦装置に関して説明したが、検索エンジンをそれに対して使用することができるいかなる種類のメディアまたはデータに、本発明を適用することができることは明らかである。したがって、本発明を、例えば、インターネット検索ツールまたは音楽データベース用の検索エンジンに関連付けて利用することができる。

30

【図面の簡単な説明】

【0052】

【図1】本発明を実施可能な1つのタイプの誘導エンジンを説明するための概念空間の図である。

【図2】2つの特定化された説明からデータを統合して、フィードバックを生成するためのソース・フィルタか、またはユーザの説明の代わりとなる新たな特定化された説明の何れかを形成する図である。

【図3】一般化された説明と特定化された説明に、別のユーザの特定化された説明を統合させて、テスト対象データ用のソース・フィルタを形成することを表す図である。

【図4】特定化された説明の特徴に対するラベルの選択を表す。

40

【図5】本発明を実施するための例示的なハードウェア環境の図である。

【図6】第一タイプの特徴-値-評点タイプのプロフィール・エンジンと使用法の図である。

【図7】第二タイプの特徴-値-評点タイプのプロフィール・エンジンと使用法の図である。

【符号の説明】

【0053】

100... 概念空間

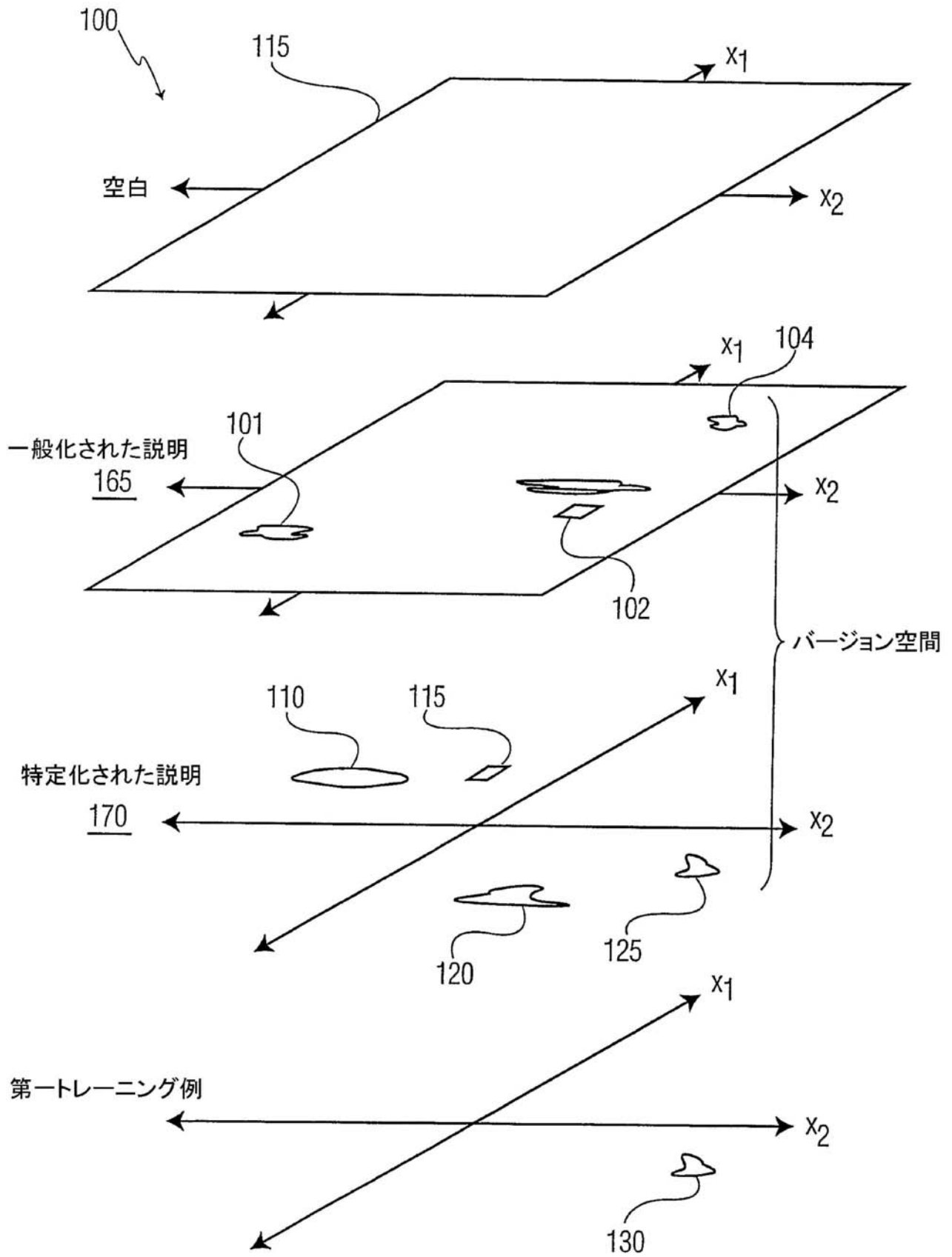
101... バージョン空間

110... 範囲

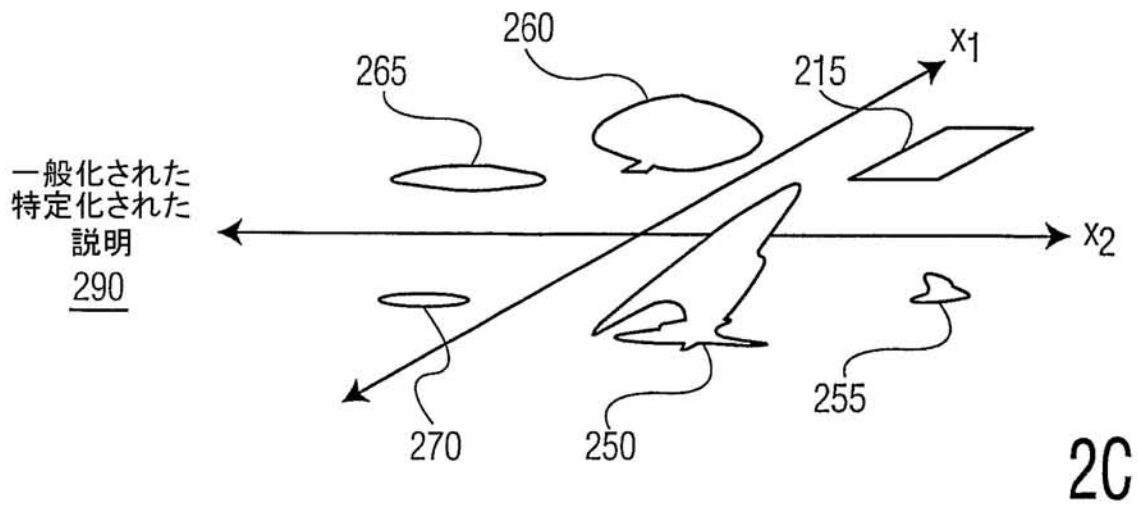
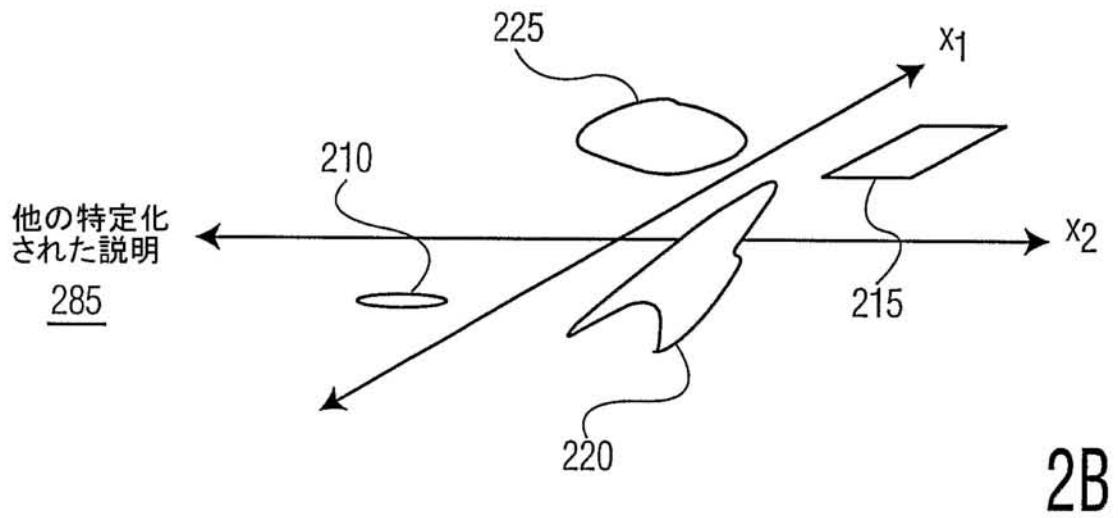
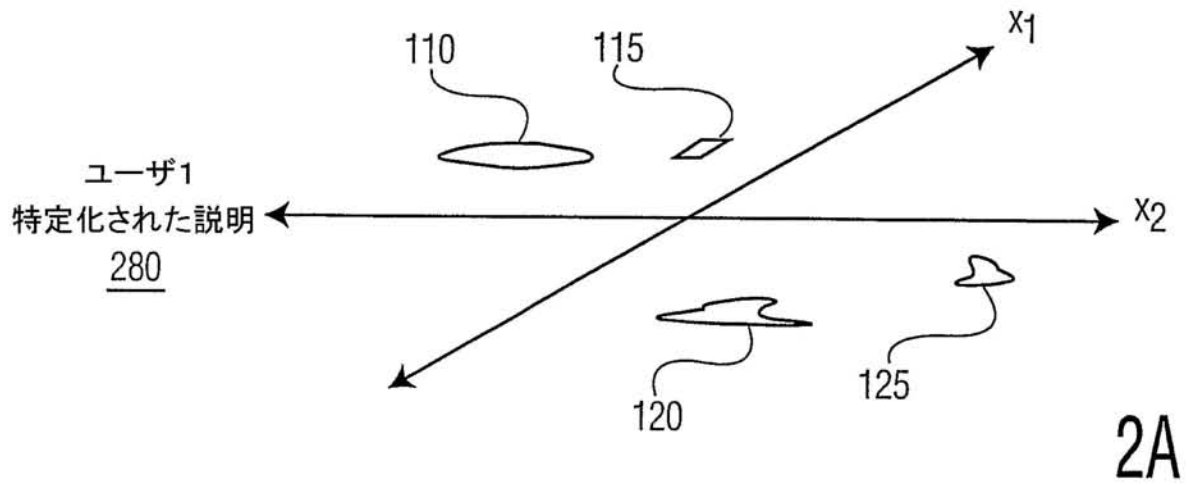
50

115... 範囲	
120... 範囲	
125... 範囲	
130... 最初の肯定的な例しか含まない単集合	
165... 最新の一般化された説明	
170... 最新の特定化された説明	
210... 説明部分	
215... 範囲	
220... 範囲	
225... 範囲	10
250... 隣接する範囲	
255... 隣接する範囲	
260... 隣接する範囲	
265... 隣接する範囲	
270... 隣接する範囲	
280... ユーザの特定化された説明	
285... 他の特定化された説明	
290... 新たな特定化された説明	
315... フィードバック用の新たなテンプレート	
330... 寸法	20
410... 遠隔制御装置	
415... 赤外線ポート	
420... メモリ・カードまたはディスク	
430... テレビのチューナ	
435... 大量格納装置	
440... コンピュータ	
445... コンピュータにリンクされたチューナ	
460... データ・リンク	
470... 映像信号	
500... ユーザ・インターフェース	30
505... フィードバック履歴ファイル	
510... 視聴履歴データベース	
520... 最新の番組	
525... フィードバック・プロフィール・データベース	
530... 内在的プロフィールのデータベース	
555... 1つ以上の選択肢	
560... フィードバック・データ	
575... 番組データベースから得ることができる選択肢がフィルタまたは分類された一覧	
580... 推薦装置	
595... プロファイラ	40
620... 推薦装置	
670... 記録	
675... 得点が添付された特徴-値の対	
680... 内在的プロフィールのデータベースのコンテンツ	
685... 推薦	

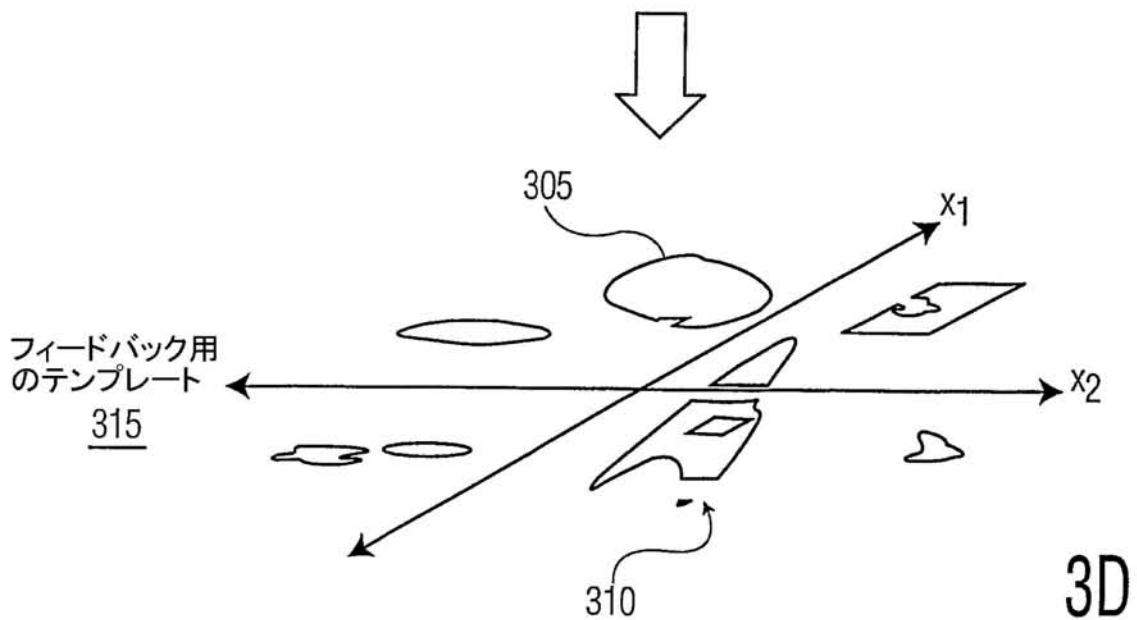
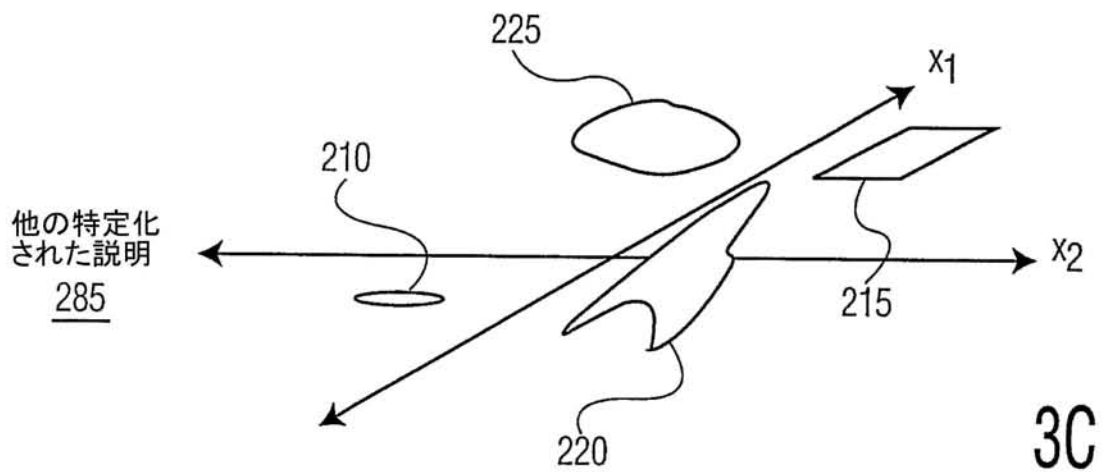
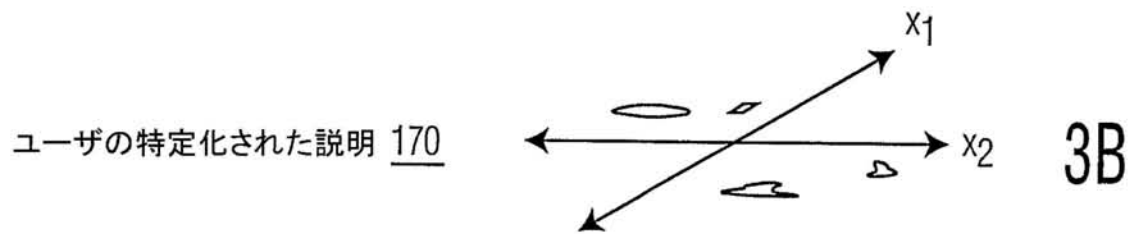
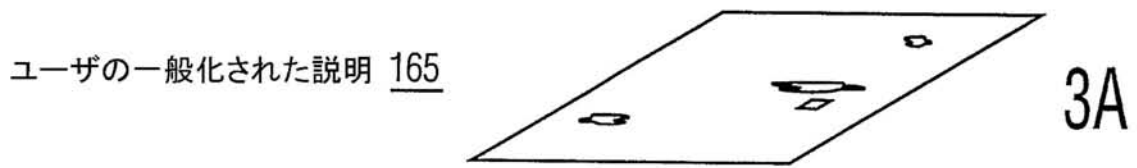
【図1】



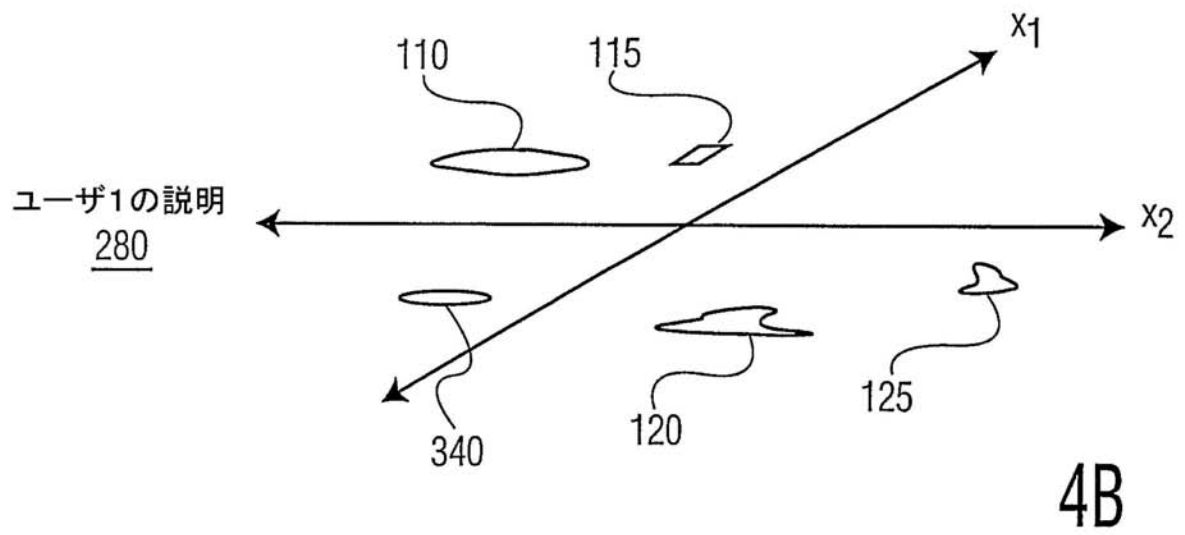
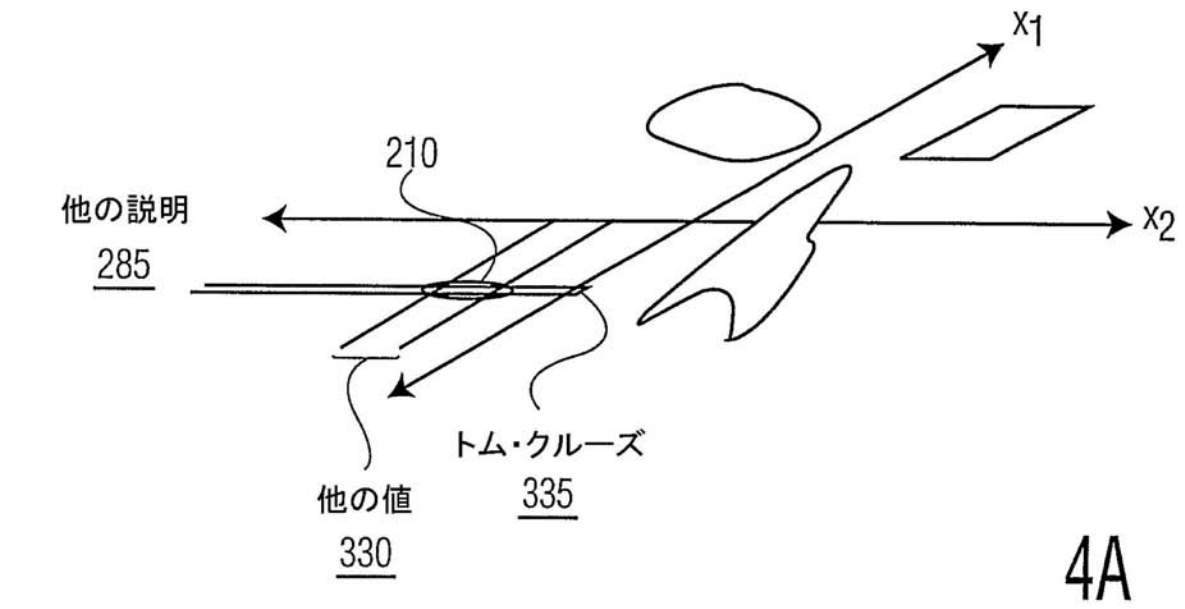
【 図 2 】



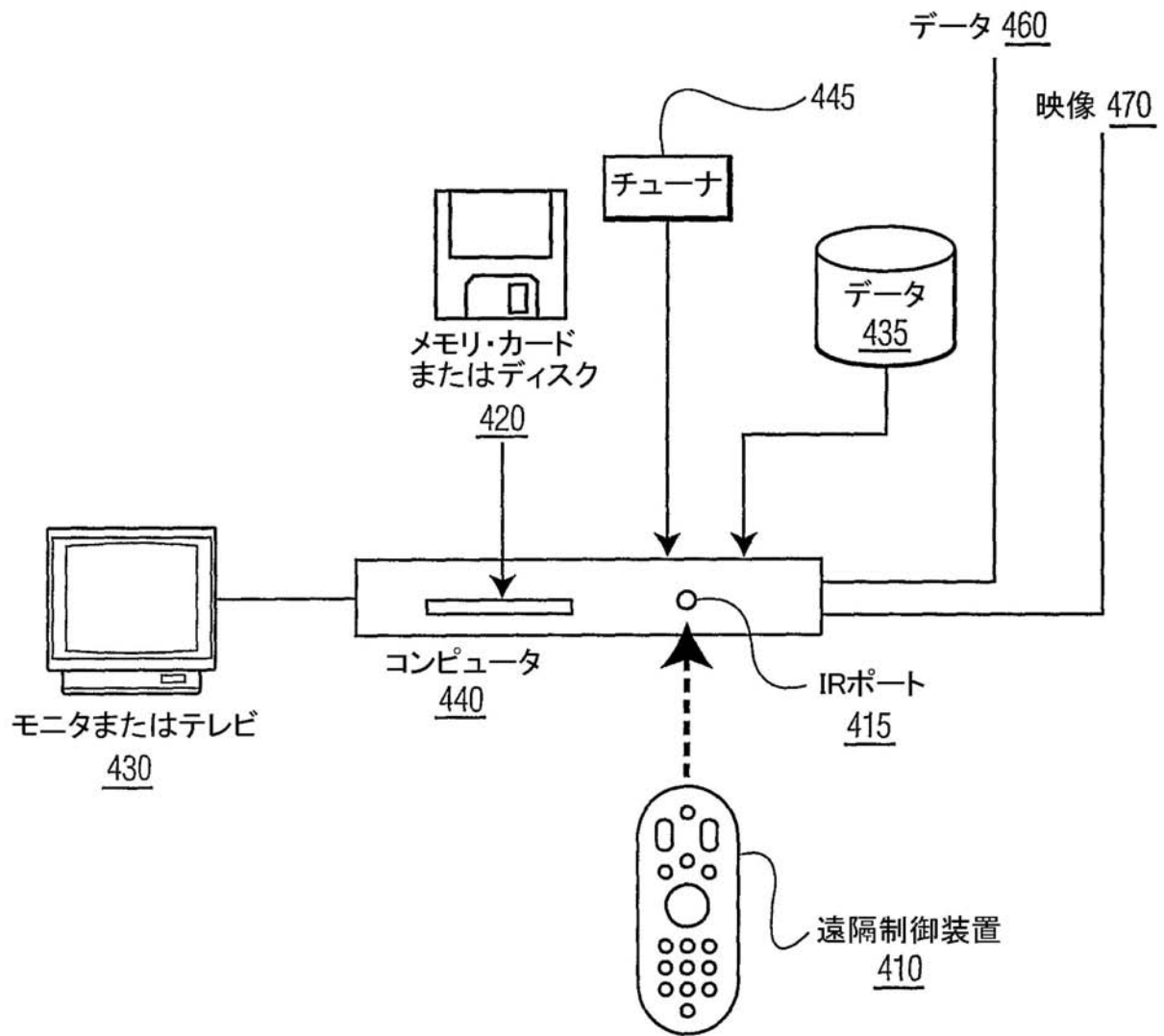
【 図 3 】



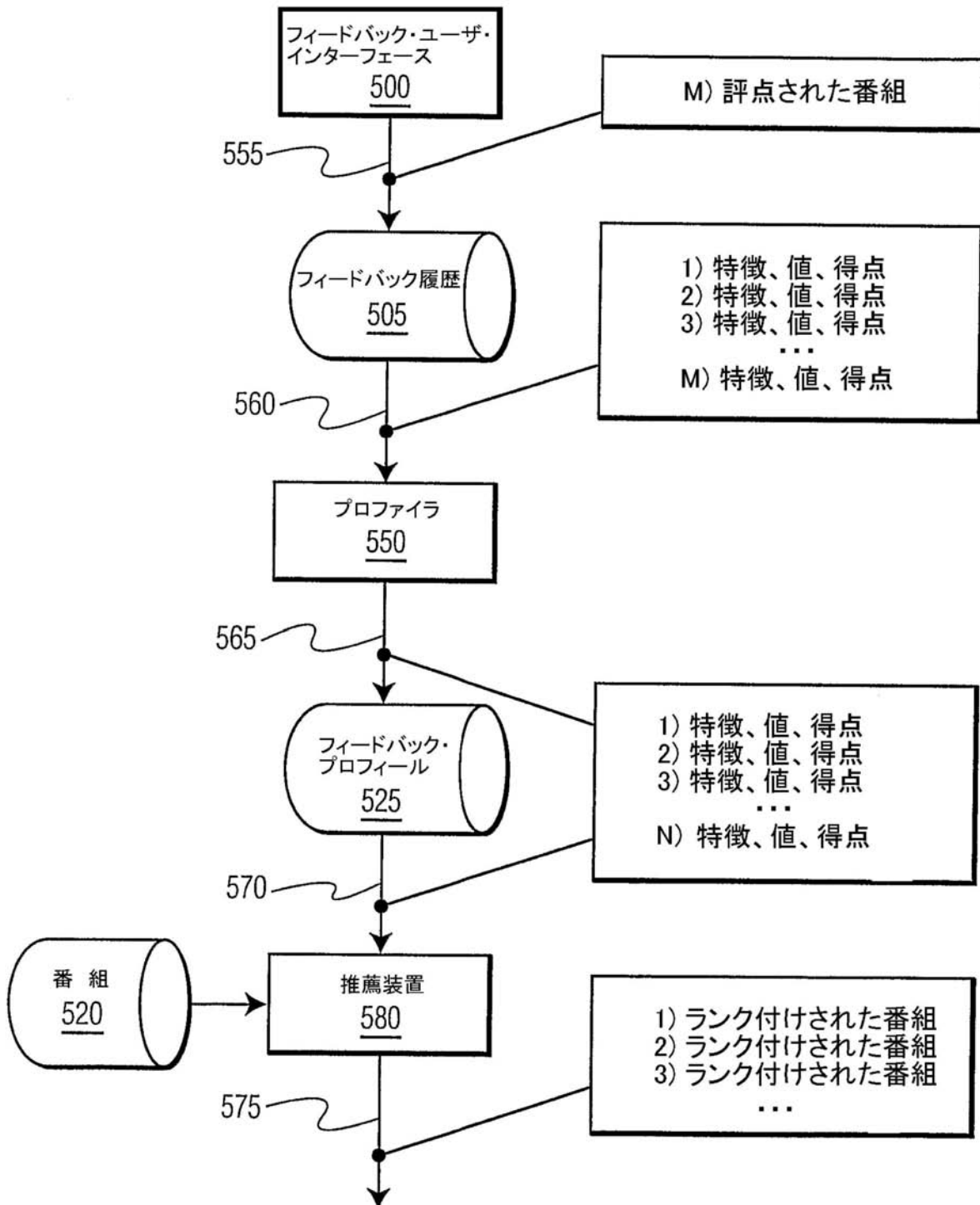
【 図 4 】



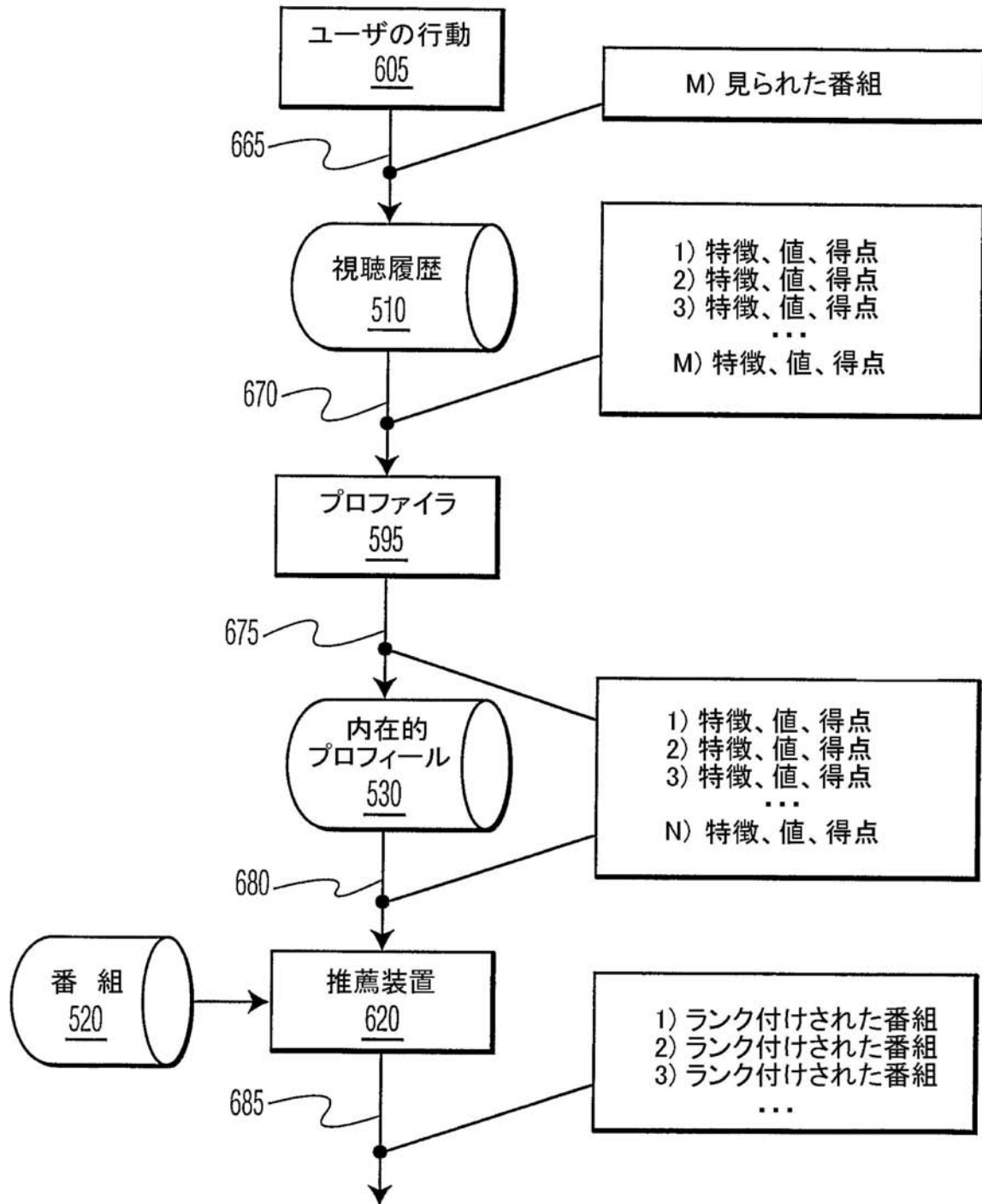
【 図 5 】



【図 6】



【 図 7 】



【国際公開パンフレット】

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

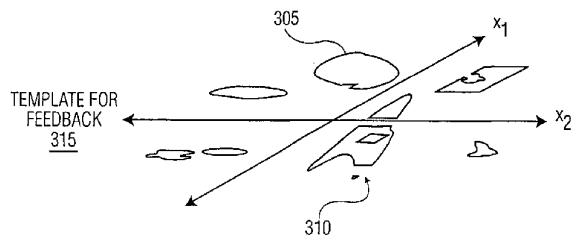
(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
10 April 2003 (10.04.2003)

PCT

(10) International Publication Number
WO 03/030528 A2

- (51) International Patent Classification: H04N 5/445 (74) Agent: GROENENDAAL, Antonius, W., M.; International Octrooibureau B.V., Prof. Holslaan 6, NL-5656 AA Eindhoven (NL).
- (21) International Application Number: PCT/IB02/03693
- (22) International Filing Date: 10 September 2002 (10.09.2002) (81) Designated States (national): CN, JP, KR.
- (25) Filing Language: English (84) Designated States (regional): European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR).
- (26) Publication Language: English
- (30) Priority Data: 09/966,414 28 September 2001 (28.09.2001) US Published: — without international search report and to be republished upon receipt of that report
- (71) Applicant: KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL). 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.
- (72) Inventors: GUTTA, Srinivas, V., R.; Prof. Holslaan 6, NL-5656 AA Eindhoven (NL); KURAPATI, Kaushal; Prof. Holslaan 6, NL-5656 AA Eindhoven (NL).

(54) Title: INDIVIDUAL RECOMMENDER DATABASE USING PROFILES OF OTHERS



(57) Abstract: A data-class recommender, such an electronic program guide that recommends television programs, avoids users getting trapped in a rut when the users select the same programming material over and over again. In an embodiment, the recommender may be programmed automatically to leverage the profile of another user to broaden the user's profile. For example, the recommender may use the target descriptions of other users in a same household of the user as a guide for broadening the user's profile. Alternatively, the household profile may be used as a filter for source material for soliciting feedback from the user. In this way, rather than simply broadening the user's range arbitrarily, guidance from other profiles, related in some way to the user, is obtained and leveraged. Note that the "relationship" can include friends, published stereotypes representing interests of the user, and others.

WO 03/030528 A2

WO 03/030528

PCT/IB02/03693

1

Individual recommender database using profiles of others

The invention relates to search engines that learn a user's preferences by observing a user's behavior and filter a large space of data based on the observed preferences. Such systems employ algorithms to infer rules from user behavior rather than require a user to enter rules explicitly. The invention relates more particularly to search engines that make
5 recommendation for an individual user based on both the user's choices and the choices of others.

Search engines are becoming increasingly important in applications in which very large databases must be used efficiently and quickly. Search engines are useful not only for searching the worldwide Web, but for store catalogs, television programming, music
10 listings, file systems, etc. In a world where the focus is shifting from information to knowledge, search engines are a huge growth area and have immense potential.

One way in which search engines are finding application is in so-called passive recommenders, which observe a user's selection behavior and make recommendations based on that behavior. This technique is used in connection with
15 electronic program guides (EPGs) for selecting television programming.

Electronic program guides (EPGs) promise to make more manageable, the task of choosing from among myriad television and other media viewing choices. Passive search engines build user-preference databases and use the preference data to make suggestions, filter current or future programming information to simplify the job of choosing, or even
20 make choices on behalf of the user. For example, the system could record a program without a specific request from the user or highlight choices that it recommends.

As mentioned above, one type of device for building the preference database is a passive one from the standpoint of the user. The user merely makes choices in the normal fashion from raw EPG data and the system gradually builds a personal preference database
25 by extracting a model of the user's behavior from the choices. It then uses the model to make predictions about what the user would prefer to watch in the future. This extraction process can follow simple algorithms, such as identifying apparent favorites by detecting repeated requests for the same item, or it can be a sophisticated machine-learning process such as a decision-tree technique with a large number of inputs (degrees of freedom). Such models,

WO 03/030528

PCT/IB02/03693

2

generally speaking, look for patterns in the user's interaction behavior (i.e., interaction with the user-interface (UI) for making selections).

One straightforward and fairly robust technique for extracting useful information from the user's pattern of watching is to generate a table of feature-value counts.

5 An example of a feature is the "time of day" and a corresponding value could be "morning." When a choice is made, the counts of the feature-values characterizing that choice are incremented. Usually, a given choice will have many feature-values. A set of negative choices may also be generated by selecting a subset of shows (optionally, at the same time) from which the choice was discriminated. Their respective feature-value counts will be
10 decremented (or a count for shows not watched incremented). These data are sent to a Bayesian predictor which uses the counts as weights to feature-counts characterizing candidates to predict the probability that a candidate will be preferred by a user. This type of profiling mechanism is described in US Patent Application Ser. No. 09/498,271, filed 2/4/2000 for BAYESIAN TV SHOW RECOMMENDER, the entirety of which is hereby
15 incorporated by reference as if fully set forth herein. A rule-based recommender in this same class of systems, which build profiles passively from observations of user behavior, is also described in the PCT application, WO 99/01984 published 1/14/99 for INTELLIGENT ELECTRONIC PROGRAM GUIDE.

Another example of the first type is MbTV, a system that learns viewers' television watching preferences by monitoring their viewing patterns. MbTV operates
20 transparently and builds a profile of a viewer's tastes. This profile is used to provide services, for example, recommending television programs the viewer might be interested in watching. MbTV learns about each of its viewer's tastes and uses what it learns to recommend upcoming programs. MbTV can help viewers schedule their television watching time by
25 alerting them to desirable upcoming programs, and with the addition of a storage device, automatically record these programs when the viewer is absent.

MbTV has a Preference Determination Engine and a Storage Management Engine. These are used to facilitate time-shifted television. MbTV can automatically record, rather than simply suggest, desirable programming. MbTV's Storage Management Engine
30 tries to insure that the storage device has the optimal contents. This process involves tracking which recorded programs have been viewed (completely or partially), and which are ignored. Viewers can "lock" recorded programs for future viewing in order to prevent deletion. The ways in which viewers handle program suggestions or recorded content provides additional feedback to MbTV's preference engine which uses this information to refine future decisions.

WO 03/030528

PCT/IB02/03693

3

MbTV will reserve a portion of the recording space to represent each "constituent interest." These "interests" may translate into different family members or could represent different taste categories. Though MbTV does not require user intervention, it is customizable by those that want to fine-tune its capabilities. Viewers can influence the "storage budget" for different types of programs. For example, a viewer might indicate that, though the children watch the majority of television in a household, no more than 25% of the recording space should be consumed by children's programs.

A second type of device is more active. It permits the user to specify likes or dislikes by grading features. These can be scoring feature-value pairs (a weight for the feature plus a value; e.g., weight = importance of feature and value the preferred or disfavored value) or some other rule-specification such as favorite programs, combinations of feature-value pairs like "I like documentaries, but not on Thursday which is the night when the gang comes over." For example, the user can indicate, through a user interface, that dramas and action movies are favored and that certain actors are disfavored. These criteria can then be applied to predict which, from among a set of programs, would be preferred by the user.

As an example of the second type of system, one EP application (EP 0854645A2), describes a system that enables a user to enter generic preferences such as a preferred program category, for example, sitcom, dramatic series, old movies, etc. The application also describes preference templates in which preference profiles can be selected, for example, one for children aged 10-12, another for teenage girls, another for airplane hobbyists, etc.

A third type of system allows users to rank programs in some fashion. For example, currently, TIVO® permits user's to give a show up to three thumbs up or up to three thumbs down. This information is similar in some ways to the second type of system, except that it permits a finer degree of resolution to the weighting given to the feature-value pairs that can be achieved and similar to the first type except the expression of user taste in this context is more explicit. (Note, this is not an admission that the Bayesian technology discussed in US Patent Application Ser. No. 09/498,271 combined with user-ranking, as in the third type of system, is in the prior art.)

A PCT application (WO 97/4924 entitled System and Method for Using Television Schedule Information) is an example of the third type. It describes a system in which a user can navigate through an electronic program guide displayed in the usual grid fashion and select various programs. At each point, he/she may be doing any of various

WO 03/030528

PCT/IB02/03693

4

described tasks, including, selecting a program for recording or viewing, scheduling a reminder to watch a program, and selecting a program to designate as a favorite. Designating a program as a favorite is for the purpose, presumably, to implement a fixed rule such as: "Always display the option of watching this show" or to implement a recurring reminder.

5 The purpose of designating favorites is not clearly described in the application. However, more importantly, for purposes of creating a preference database, when the user selects a program to designate as a favorite, she/he may be provided with the option of indicating the reason it is a favorite. The reason is indicated in the same fashion as other explicit criteria: by defining generic preferences.

10 The first type of system has the advantage of being easier on the user since the user does not have to provide any explicit data. The user need merely interact with the system. For any of the various machine-learning or predictive methods to be effective, a substantial history of interaction must be available to build a useful preference database. The second and third types have the advantage of providing explicit preference information. The
15 second is reliable, but not perfect as a user may have a hard time abstracting his own preferences to the point of being able to decide which criteria are good discriminators and what weight to give them. The third does not burden the user and probably provides the best quality of information, but can be a burden to generate and still may not contain all the information that can be obtained with the second and also may require information on many
20 shows like the first.

One of the problems with prior art techniques for building preference
databases manifests when a user repeatedly watches the same program. A large percentage of
the user's choices are made up of too small a set of data and rules extracted from these
choices end up defining an overly narrow range of recommendations. The problem is akin to
25 falling into a rut. Another problem with prior art techniques is that they do not permit the easy sharing of implicit profiles among users. If a user likes the recommendations of a friend, there is no good way for the user to obtain some or all parts of his/her friend's profile and combine it in some way with his/her own.

The invention provides mechanisms to expand the choices provided by a
30 user's preference profile based on the preferences of others, particularly those of users in the same household. Various types of mechanisms for generating and refining a selection engine based on positive and/or negative examples are known. One, called a version space algorithm, saves two descriptions of all the possible choices available in a database (i.e., the "choice space": (1) a general description that is the broadest description of the choice space

WO 03/030528

PCT/IB02/03693

5

excludes all negative choices and (2) a specialized description that is the narrowest description that embraces all positive examples in the choice space. Each time a negative or positive example is provided, it is used to alter the specialized or generalized description accordingly. The algorithm and further details on the version space algorithm is described in
5 US Patent Application Ser. No. 09/794,445 entitled "Television Programming Recommendations Through Generalizations And Specialization Of Program Content," which is hereby incorporated by reference as if fully set forth herein in its entirety.

In the sphere of television program selection, the generalized description indicates all the possible programming choices that a user might be interested in. The
10 specialized description indicates all the possible programming descriptions the user is clearly interested in. The range of descriptions between the generalized and specialized descriptions can be great. Also, the generalized description can be too liberal to reduce a large set of selections to a reasonable number and the specific description can be overly narrow for being trapped by a narrow range of examples.

The prior art has offered other ways to bump a user out of this mess. One is to select program content at random from the large space defined by the generalized description and ask the user to rank them. But this can lead to pretty stupid exercises. For example, suppose the only examples provided are English-language examples. The user has given no
15 negative examples of content in the space of non-English descriptions. But most users are likely to be disinclined to expand their language horizons by watching television. Thus, a random selector would grab examples outside the English language space and ask the user to rank them only to get criteria that are marginally useful. That is, did the user not like it because it was about cars or because it was in the Spanish-language? A user would be quick to become bored if he were asked to rank too many irrelevant choices. It would be better to
20 pull examples from a narrower description than the user's generalized description. According to the invention, this may be done by leveraging the specialized description or descriptions of others who are similar to the user according to some criterion, for example, users in the same household.

In one embodiment, a generalized-specialized description is defined that
30 embraces the entire space of specialized descriptions of one or more other persons selected by the user. This generalized-specialized description is used as a source filter for generating test-samples with respect to which the user's positive and negative feedback is solicited. In another embodiment, a group is defined automatically, such as all the users in a household, and a new specialized description generated that is the narrowest to embrace the spaces

WO 03/030528

PCT/IB02/03693

6

defined by all the specialized descriptions. Test-samples are similarly derived from the new specialized space.

5 In a refinement of both of the above embodiments, priority is given to test-samples that discriminate ambiguous dimensions in the user's specialized description. That is samples from the generalized-specialized description that conform to the user's specialized description already are avoided and samples that are outside that description are favored. The latter samples clearly have higher discriminating power in the dimensions along which the user's specialized description is confluent with that of the generalized-specialized description.

10 Another refinement of the above approaches is to use the user's generalized description to specialize the generalized-specialized description. Because the generalized description is the storehouse of what the user doesn't like, it can be used as a filter to filter the space of the generalized-specialized description.

15 In another embodiment, classes of users are defined and, in a manner akin to collaborative filtering, the user's specialized description is generalized to embrace the space of the specialized descriptions of archetypal users. For example, a service provider may generate specialized descriptions for stereotypes such as: "sports fanatic," "blood and guts," "history geek," "mawkishly sentimental," "science lover," and "fantasy lover."

20 In yet another embodiment, rather than use other specialized descriptions to create a source for feedback to refine the user's descriptions, a new specialized description is created leveraging other specialized descriptions. In other words, the generalized-specialized description is substituted for the specialized description of the user.

25 In a user interface supporting an embodiment in which the user's specialized description is substituted for the generalized-specialized description, the user may be asked to try a stereotype out for a period of time. The old specialized description may be retrieved if the user did not like the result. Optionally, the user may retain the benefit of feedback obtained while the stereotypic description was applied to generalize the user's specialized description.

30 The invention can be extended to other types of induction engines. For example, neural networks can be trained on predictions from other networks to generalize their predictions of likes and dislikes. Decision trees can be expanded by known techniques such as by adding samples generated by another decision tree or more directly by sharing branches from another decision tree. Other types of machine learning, even ones as yet unknown, can also use the basic ideas behind the invention and should be within the

WO 03/030528

PCT/IB02/03693

7

competence of one skilled in the art in combination with the teachings in the present application.

The invention will be described in connection with certain preferred embodiments, with reference to the following illustrative figures so that it may be more fully understood. With reference to the figures, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only, and are presented in the cause of providing what is believed to be the most useful and readily understood description of the principles and conceptual aspects of the invention. In this regard, no attempt is made to show structural details of the invention in more detail than is necessary for a fundamental understanding of the invention, the description taken with the drawings making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

Fig. 1 is an illustration of a concept space for purposes of describing one type of induction engine in which the present invention may be implemented.

Figs. 2A-2C are illustrations of the aggregation of data from two specialized descriptions to form either a source filter for generating feedback or a new specialized description to be substituted for that of a user.

Figs. 3A-3D are illustrations representing the aggregation of generalized and specialized descriptions with the specialized description of another user to form a source filter for test target-data.

Figs. 4A and 4B illustrate selection of a label for a specialized description feature.

Fig. 5 is an illustration of an example hardware environment for implementing the invention.

Fig. 6 is an illustration of a first type of feature-value-score type of profile engine and use.

Fig. 7 is an illustration of a second type of feature-value-score type of profile engine and use.

Referring to Fig. 1, a concept space 100 is defined in terms of a description formalism. For example, Fig. 1 is suggestive of a frame-based data structure or representation

WO 03/030528

PCT/IB02/03693

8

language using a Venn-type representation for the values in each frame-slot. For purposes of discussion, the large number of slots in the frame-based structure are represented as two axes, x_1 and x_2 which represent descriptor components, such as a slot in a frame-based structure. It is to be understood that the slots chosen may represent any parameters and the diagram is not intended to suggest that they are independent or that there is any limit on their number. For example, axis X_1 could represent type of television show (comedy, drama, horror, sports, etc.) and x_2 could represent actors (Tom Cruise, Shelly Duvall, Robert Wagner, etc.) For purposes of discussion, it can be imagined that there are many different descriptor components, each of which may take on one or more values or ranges of values and each of which may or may not be dependent of another descriptor component.

A universe of possible descriptions (the concept space 100) is limited only by the inherent bias of the formalism. Here, every possible description is contained in a null generalized description 115 at the highest level of a concept space. Before any learning has occurred, this singleton generalized description 115 embraces every possible example. At the lowest level of the concept space is a singleton which embraces only the first positive example 130 provided by a user.

After training for a period of time with positive and negative examples, for example using the version space algorithm described in the application incorporated by reference above, a most recent specialized description 170 is broadened so that it is the narrowest set of descriptions that encompasses all positive examples. By definition, it excludes all negative examples. Also, after training, a current generalized description 165 has been derived from the null generalized description 115 that is the broadest set of possible descriptions that does not contain any of the negative examples. By definition, this contains all positive examples.

Selections from the space of selections defined by the current specialized description 170 include only selections that are similar to previous positive examples. Thus, if recommendations are derived from the current specialized description 170, the recommendations will be too narrow and the user will be stuck in his/her rut for having given positive feedback on too narrow a set of examples. In such a case, the user may also have too broad a generalized description, so the generalized description may be too broad a space to expand into. There is a space, called the version space 101 lying between these extremes which defines the possible descriptions for subject matter the user might like with certainty increasing as one moves from the generalized description toward the specialized description.

WO 03/030528

PCT/IB02/03693

9

Referring now to Figs. 2A-2C, a new specialized description 290 is derived from the union of the user's specialized description 280 with another specialized description 285. The latter may be, for example, a stereotype description or one of another user. Here the user's set which is the union of domains 110, 115, 120, and 125 is combined with the other set, which is the union of domains 210, 215, 220, and 225. The result is the set defined by the union of contiguous domains 250, 255, 260, 265, 270, and 275 shown in Fig. 2C. More precisely, the new description is the user's specialized description 280 generalized so as not to exclude subject matter that is embraced by the other specialized description 285. Note that, preferably, the generalized-specialized domain includes the multiple other specialized domains of other users in a same household as the user. It has been found that expanding in a manner consistent with the other household users provides better predictions than a user's own profile.

The use of additional user profiles to expand a profile that is mired in a rut can be made selectable by the user. The user may be provided with the option of selecting a group of user profiles, a stereotyped profile, or one or more specific profiles to be used to expand the user's options. The other profiles may be used to modify the user's profile permanently or simply to expand the range of selections on a use-by-use basis. Another possibility is for the learning engine to detect when a user's profile has fallen into a rut and take corrective action, such as by adding the specialized description of all members of a household. This can be determined in various ways according to the type of profile. For example in a feature-value-score-type profile, a profile with only a small number of feature-value-score records could be identified as in a rut. In a concept space, a specialized description that is highly specialized would indicate the profile is in a rut. Note that it may be appropriate to distinguish household members of the same age and only share descriptions when the members are in a similar age category.

As is known in the prior art, a system can solicit feedback on new examples selected at random. However, such a strategy can be impractical because it may include material for which negative feedback has been provided and could just include too large a space of possible subject matter. There is a high likelihood that mostly negative examples will be found and the user would likely become frustrated and lose interest. Alternatively, the current generalized description 165 could be used as a filter for new examples. However, the current generalized description 165 may still define too large a space of possibilities to be practical.

WO 03/030528

PCT/IB02/03693

10

One approach to this problem is to use the specialized description of another user as a filter for soliciting feedback. The system may use the specialized description of another user's profile as a filter for selecting new material and request the user's feedback on that new material. Referring to Figs. 3A-3D, it is preferred that the material for which the user has already given feedback be excluded from test-examples. Thus, the corresponding portions in the user's generalized description 165 and the user's specialized description 170 may be removed from the other specialized description 285 to provide a new template for feedback 315. Although only one other specialized description 170 is shown in the figures, it is clear that the union of any number of specialized descriptions could also be used to generate a template for feedback.

One important issue relating to permitting a user to use the profiles of others to enhance his/her own profile is giving the user some sense of control over the process. Probably the dominant concern here is making it clear to the user what s/he may do. In some cases, the leveraging of other profiles may be done transparently. For example, rather than relying solely on a user's individual profile, a recommender may include recommendations that are derived from the profiles of other users in the same household as the user. This can be done part of the time or all of the time. Of course, whenever feedback is obtained, it may be used to refine the profile of the individual user.

Although the above discussion employed figurative terms and drawings suggested by version space algorithms, the invention is applicable to other types of recommender systems as well. Suppose a first user likes the examples recommended by the profile of another user. One way to permit the first user to modify his own profile using the other user's profile is to use the other user's profile to generate suggested shows using the other user's profile and permit the first user to give feedback on them. This could be done without their being any compatibility between the recommendation engines.

Another strategy for expanding a user's profile is to substitute the generalized description of another user for the generalized description of the user.

Referring to Fig. 5, an example of a hardware environment that may support the present invention includes a computer 440 equipped to receive the video signal 470 and control the channel-changing function, and to allow a user to select channels through a tuner 445 linked to the computer 440 rather than through the television's tuner 430. The user can then select the program to be viewed by highlighting a desired selection from the displayed program schedule using the remote control 410 to control the computer. The computer 440 has a data link 460 through which it can receive updated program schedule data. This could

WO 03/030528

PCT/IB02/03693

11

be a telephone line connectable to an Internet service provider or some other suitable data connection. The computer 440 has a mass storage device 435, for example a hard disk, to store program schedule information, program applications and upgrades, and other information. Information about the user's preferences and other data can be uploaded into the computer 440 via removable media such as a memory card or disk 420.

Note that many substitutions are possible in the above example hardware environment and all can be used in connection with the invention. The mass storage can be replaced by volatile memory or non-volatile memory. The data can be stored locally or remotely. In fact, the entire computer 440 could be replaced with a server operating offsite through a link. Rather than using a remote control to send commands to the computer 440 through an infrared port 415, the controller could send commands through a data channel 460 which could be separate from, or the same as, the physical channel carrying the video. The video 470 or other content can be carried by a cable, RF, or any other broadband physical channel or obtained from a mass storage or removable storage medium. It could be carried by a switched physical channel such as a phone line or a virtually switched channel such as ATM or other network suitable for synchronous data communication. Content could be asynchronous and tolerant of dropouts so that present-day IP networks could be used. Further, the content of the line through which programming content is received could be audio, chat conversation data, web sites, or any other kind of content for which a variety of selections are possible. The program guide data can be received through channels other than the separate data link 460. For example, program guide information can be received through the same physical channel as the video or other content. It could even be provided through removable data storage media such as memory card or disk 420. The remote control 410 can be replaced by a keyboard, voice command interface, 3D-mouse, joystick, or any other suitable input device. Selections can be made by moving a highlighting indicator, identifying a selection symbolically (e.g., by a name or number), or making selections in batch form through a data transmission or via removable media. In the latter case, one or more selections may be stored in some form and transmitted to the computer 440, bypassing the display 170 altogether. For example, batch data could come from a portable storage device (e.g. a personal digital assistant, memory card, or smart card). Such a device could have many preferences stored on it for use in various environments so as to customize the computer equipment to be used.

Some types of profiling mechanisms permit their internal target descriptions to be displayed as abstractions. For example, it would be possible in a frame-based data

WO 03/030528

PCT/IB02/03693

12

structure to actually allow one user to inspect another user's profile by associating titles with the different slots. Although the influence of a choice in any one slot can influence allowed choices in other slots, because the slots are not independent, it is not necessarily a straightforward task to present to a user a meaningful view of how a profile is constructed.

- 5 For example, a user's profile may contain a specialized description that suggests the actor Tom Cruise is favored by the user. But the examples for which positive feedback was given are restricted to action-type movies. Thus, it cannot be said that the user likes Tom Cruise. It may be that the user only likes Tom Cruise in certain types of movies. The above example is simple. The real examples could be very complex and therefore make it difficult to present to
- 10 user. The interface would have to show all the linked slots with any slot of interest thereby defining a multiple-parameter space. But consider that the goal is not to be 100% precise. The goal may be simply to permit the user to borrow only certain aspects of another user's profile and characterizing that aspect may not have to be so complete. The system could offer to modify a user's profile based on a particular slot that is coupled with many other slots by
- 15 tagging the modification based on the values in only one slot. Thus, if the system indicated to a first that a second user's profile showed a marked preference for Tom Cruise, the first user, in accepting a modification to his/her own profile based on that preference, could expand his/her profile so that it recommended Tom Cruise examples coupled with all the attendant caveats implicit in the second user's profile. In other words, in the example given, the first
- 20 user would be asked if s/he wants Tom Cruise and s/he would get Tom Cruise, but only Tom Cruise in action movies.

- Determining labels such as "Tom Cruise" for the features of a user's profile, in a frame-based data structure conditioned under the version space algorithm, could be identified by selecting a value (e.g., "Tom Cruise" that appears in combination many times
- 25 with values in other slots. In other words, there is a high incidence of that slot-value in the specialized description. This mechanism for permitting a user to control the porting of description information from one profile to another is illustrated in Figs. 4A and 4B. Here, a user's description, which could be, for example, the user's specialized description, is scanned and various portions of it labeled according to a dominant feature. Shown in the figure is the
- 30 labeling of a portion 210 as "Tom Cruise." Figuratively speaking, one dimension of the data structure x_1 may correspond to actor. The other dimension, x_2 , may be considered to correspond to other parameters such as type of movie or any other. The value "Tom Cruise" has been selected in association with multiple values of other parameters so it may be inferred that it is an important feature-value.

WO 03/030528

PCT/IB02/03693

13

Note that although the portion 210 of the description is shown as a contiguous closed space, as are the other portions in the other figures, which suggests contiguous ranges, such a feature may or may not represent how data is represented in a target description. In a frame-based model, each feature or slot may take on discrete values and there may be no relationship between adjacent features such that data sets would tend to form closed spaces such as 210. This is merely an abstraction borrowed for purposes of discussion. The only aspect of the closed space is that its length in the dimension indicated at 330 is suggestive of the fact that the value "Tom Cruise" is associated with multiple values of the other feature along dimension x_2 suggesting its importance.

In other types of data structures, mechanisms for labeling portions of a profile would be readily identified. For example, in systems that store feature-value pairs labeling an important feature and porting that feature to another profile is even easier. Referring to Fig. 6, in such a system the user provides feedback to rank a choice as liked or disliked and, optionally, includes a degree of like or dislike. For example, a system may use a score from 1-7 with 4 being neutral, 1-3 representing degrees of dislike and 5-7 representing degrees of liking. A user interface (UI) 500 is used to list programs and accept the feedback information. Alternatively, the UI 500 may be a simple prompt that requests the user to give feedback on a program when the program either ends or when the user switches away from the program. Preferably, the prompt-type would be subject to a preference set that would allow the user to override the prompting in some or all situations if desired.

The information generated by each instance of the feedback UI 500 is one or more choices (shows, if it is a television database) 555 with a score associated with the choice. This is used to charge a feedback history file 505 which can contain a large number of such entries. The feedback data 560 may then be applied to a profiler 550. Alternatively the data can be stored in reduced form by reducing it in a profiler 550 first and then storing in a feedback profile database 525. The reduction may be a set of feature-value pairs 465, each with a ranking as described in 09/498,271, filed 2/4/2000 for BAYESIAN TV SHOW RECOMMENDER. A given choice may give rise to a number (M) feature value pairs 565 with corresponding scores. Preferably, the user rates programs that are both liked and disliked so that both positive and negative feedback are obtained. If only positive feedback is acquired, say because feedback is only provided for programs selected for viewing, then the negative factors may not populate the database. This can be improved then, by having the system generate a set of negative choices by selecting a subset of shows available at the same time the choice was made. Preferably, as stated, the user provides a balance of positive and

WO 03/030528

PCT/IB02/03693

14

negative feedback and the automatic sampling of negative choices is not required. Their respective feature-value counts would be decremented. This data stored over many choices may be stored in the feedback profile 525 database. The entire body of N records 555 is then available when the recommender 580 makes recommendations based on a list of candidates
5 derived from a show database 520. The end result of this process is a filtered or sorted list 575 of choices available from the show database 520. The recommender may be a Bayesian filter or any other predictor.

Referring to Fig. 7, a very similar process as in Fig. 6 may be used to generate a feature-value pair profile database. This predictor is of the first type described in the
10 background section. Here, a user's selection of a program choice is inferred to indicate a positive score for a program choice. The result of a given choice by a user is a particular program 665 optionally with an attending score. This result can also include a score which may be inferred from the way the user responded. If the user watched the program to completion, the score may be high and if watched for only a short time, the score could be
15 negative. If the program were watched for a period between these two, the score could be a middle magnitude. Alternatively, a watched program could receive a positive score and a random sample of unwatched programs (optionally, at the same time) a negative score.

The view history database 510 stores the shows and scores. The records 670 are supplied to a profiler 595 which generates feature-value pairs with attending scores 675,
20 which may be stored in an implicit profile database 530. The contents 680 of the implicit profile database 530 are then available to a recommender 620 which combines them with data from current shows 520 to generate recommendations 685.

In this type of profiler, the lack of coupling of features makes uncomplicated the problem of labeling the parts of the data that may be ported from one profile to another.
25 Thus, the feature "actor" and value "Tom Cruise" would be easy to identify as standing out in a target profile. This is because that feature-value pair would have a high score associated with it. A user could be offered the option of selecting that aspect of another user's profile for porting over into his/her profile. The result would be an adjustment of the score associated with the corresponding feature-value pair in the user's profile.

30 Combining the feature-value-score type data to broaden a user whose profile is in a rut would be a matter of, in the rutted user's profile, raising the scores of feature-value pairs that have high scores in the non-rutted user's databases. Again, a user interface could be generated to allow the rutted user to select the feature-values to be modified. Alternatively, the user could permit it to be done blindly. Yet another alternative to allow the change to be

WO 03/030528

PCT/IB02/03693

15

done only temporarily to try the change out. Another way to handle the falling-into-a-rut problem is to adjust any very strong scores associated with a user's profile. This could be done selectively by the user. The user interface could indicate to the user what feature values have very strong scores (either positive or negative) and permit the user to modify them.

5 It will be evident to those skilled in the art that the invention is not limited to the details of the foregoing illustrative embodiments, and that the present invention may be embodied in other specific forms without departing from the spirit or essential attributes thereof. The present embodiments are therefore to be considered in all respects as illustrative and not restrictive, the scope of the invention being indicated by the appended claims rather
10 than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced therein.

For example, although the invention was discussed with reference to a television recommender, it is clear it is applicable to any kind of media or data for which a search engine might be used. Thus, for example, the invention could be used in the context of
15 an Internet search tool, or search engine for a music database.

WO 03/030528

PCT/IB02/03693

16

CLAIMS:

1. A method of modifying a first user's user profile for a data-class recommender, comprising the steps of:
receiving feedback from a first user scoring examples falling into various data-classes;
5 refining said first user's user profile responsively to said feedback;
modifying said first user's user profile responsively to data from a second user's user profile;
said step of modifying including modifying such that a frequency of recommendations of at least one data-class is increased without decreasing a frequency of
10 recommendations of any other data-classes, whereby said first user's user profile is expanded in scope according to preferences stored in said second user's user profile.
2. A method as in claim 1, wherein said first user's user profile includes a specialized target description of favored data-classes and said step of modifying includes
15 generalizing said specialized target description such that it encompasses at least one specialized target description of said second user's user profile.
3. A method as in claim 2, wherein said step of modifying includes substituting
at least a union of specialized descriptions of said first user's user profile and said second
20 user's user profile for said specialized description of said first user's user profile.
4. A method as in claim 2, wherein said step of generalizing includes substituting
at least a union of specialized descriptions of said first user's user profile and said second
user's user profile for said specialized description of said first user's user profile.
25
5. A method as in claim 1, comprising the steps of:
selecting test-data for revising said first user's user profile responsively to data
from at least a second user's user profile; and

WO 03/030528

PCT/IB02/03693

17

requesting feedback on said test-data from said first user and modifying said first user's user profile responsively to said feedback.

6. A method as in claim 5, wherein said step of selecting includes selecting only test-data for which feedback incorporated in said first user's profile increases a discriminating power of said first user's user profile.

7. A method as in claim 5, wherein said selecting includes selecting primarily test-data for which said first user's user profile is insufficient for said recommender to determine whether said test-data would be favored or disfavored.

8. A method as in claim 5, wherein said step of selecting includes filtering a universe of data choices through a specialized description of a concept space.

9. A data-class recommender, comprising:
a learning engine (440, 550);
a user interface device (410, 420, 500, 605) connectable to said learning engine;
said learning engine being connectable to a data source (435, 445, 520)
containing descriptions of data selections;
said learning engine being programmed to receive, through said user interface, feedback from a first user evaluating said data selections and to progressively generate a description of data selections that are favored and disfavored by said first user, thereby generating a first user profile (525, 530);
said learning engine being further programmed to generate recommendations of data selections for said first user responsively to said first user profile;
said learning engine being further programmed to selectively generate recommendations of data selections for said first user responsively to said first user profile and at least a second user profile of a second user.

10. A recommender as in claim 9, wherein said learning engine is programmed such that said first user profile includes a narrow description defining target data selections and a broad description defining non-target data selections, the recommendations being derived from a space of selections lying between said broad and narrow descriptions.

WO 03/030528

PCT/IB02/03693

18

11. A recommender as in claim 9, wherein said learning engine is programmed such that said first user profile includes at least a narrow description defining target data selections and said learning engine is further programmed to compare a level of narrowness in said narrow description to a threshold such that said first user profile results in
5 recommendations embracing a range of target data that is narrower than said threshold and said learning engine is further programmed to selectively generate recommendations of data selections for said first user responsively to said first user profile and said at least a second user profile responsively to a result of so-comparing said level with said threshold.

WO 03/030528

PCT/IB02/03693

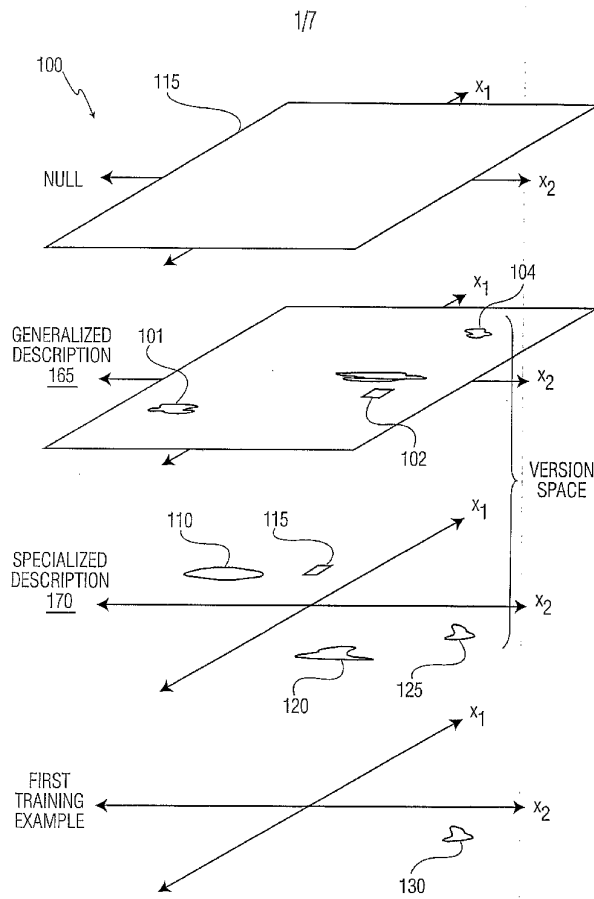
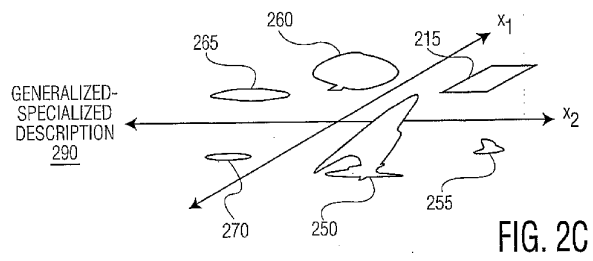
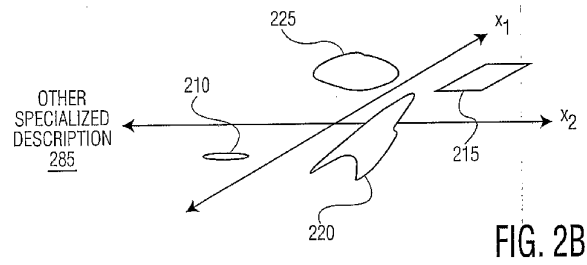
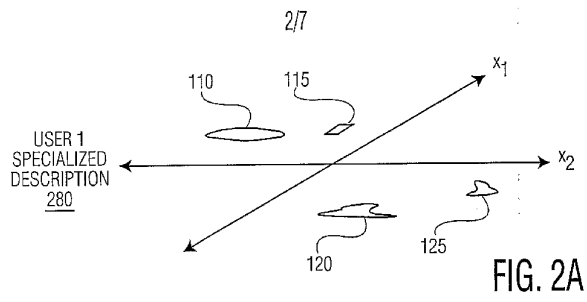


FIG. 1

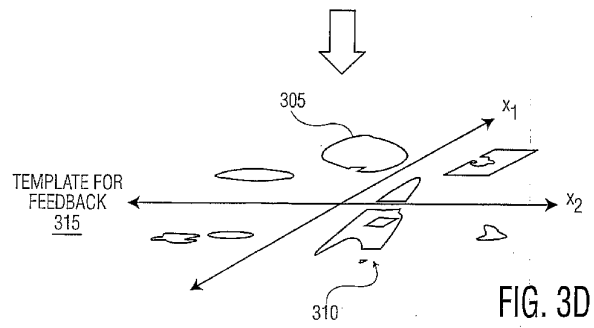
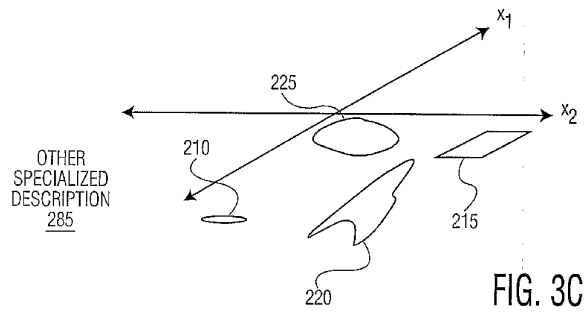
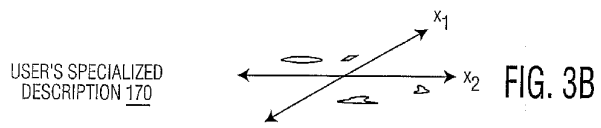
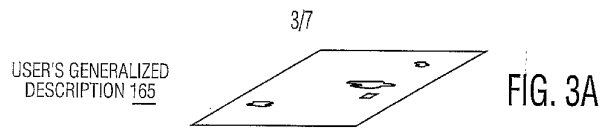
WO 03/030528

PCT/IB02/03693



WO 03/030528

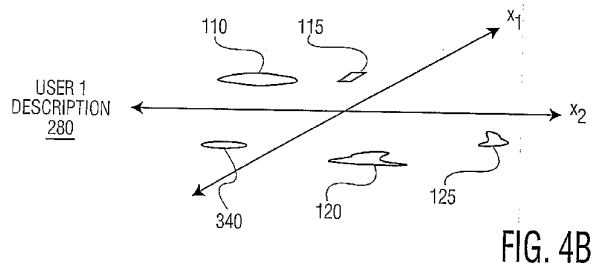
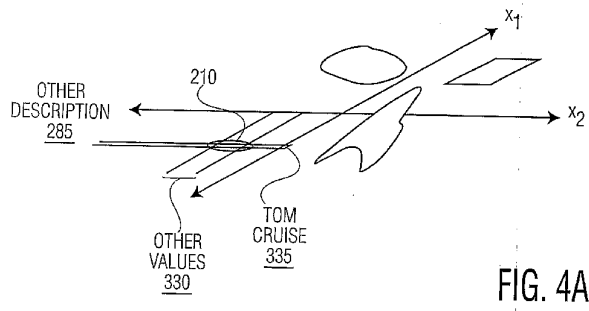
PCT/IB02/03693



WO 03/030528

PCT/IB02/03693

4/7



WO 03/030528

PCT/IB02/03693

5/7

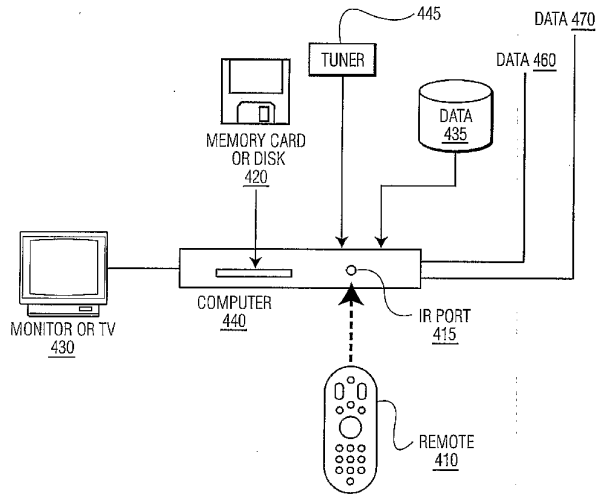


FIG. 5

6/7

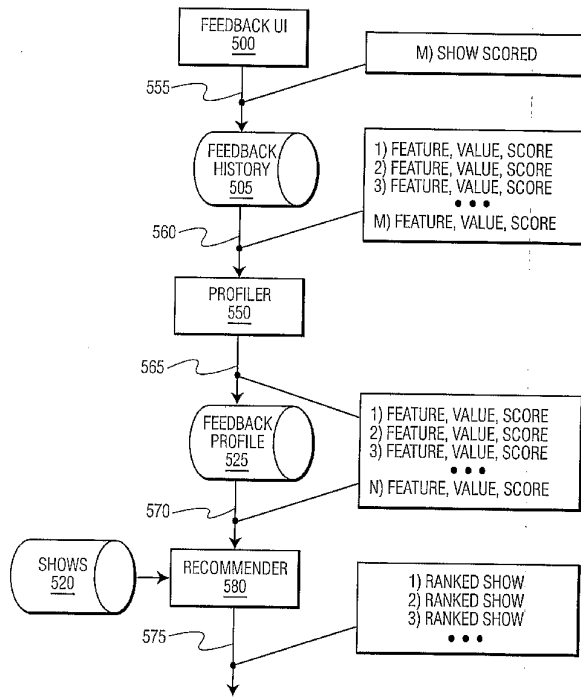


FIG. 6

WO 03/030528

PCT/IB02/03693

7/7

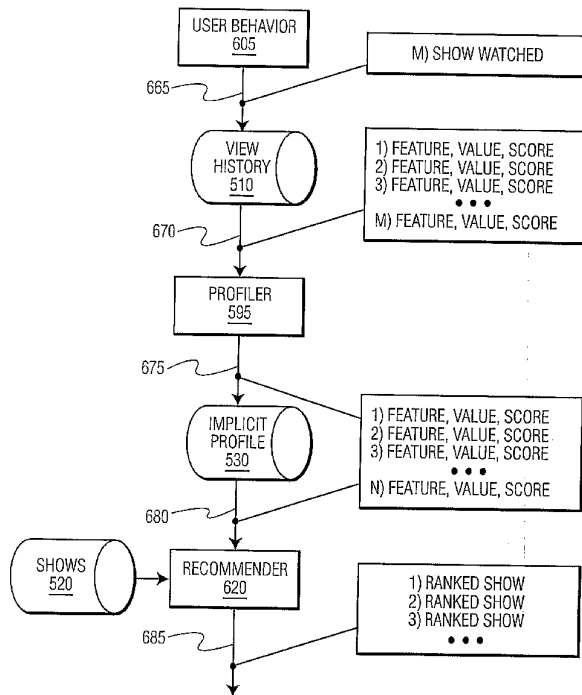


FIG. 7

【国際公開パンフレット（コレクトバージョン）】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
10 April 2003 (10.04.2003)

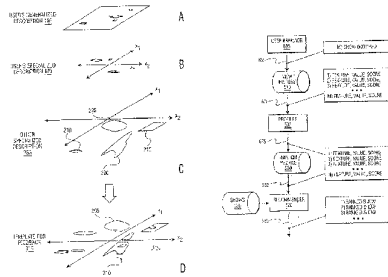
PCT

(10) International Publication Number
WO 03/030528 A3

- (51) International Patent Classification: G06F 17/30, H04N 7/16 (74) Agent: GROENENDAAL, Antonius, W., M.; International Octrooibureau B.V., Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
- (21) International Application Number: PCT/IB02/03693 (81) Designated States (national): CN, JP, KR.
- (22) International Filing Date: 10 September 2002 (10.09.2002) (84) Designated States (regional): European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR).
- (25) Filing Language: English (26) Publication Language: English
- (30) Priority Data: 09/966,414 28 September 2001 (28.09.2001) US Published: — with international search report
- (71) Applicant: KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL). (88) Date of publication of the international search report: 2 October 2003
- (72) Inventors: GUTTA, Sriulvas, V., R.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL); KURAPATI, Kaushal; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

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.

(54) Title: PERSONALIZED RECOMMENDER DATABASE USING PROFILES OF OTHERS



(57) Abstract: A data-class recommender, such as an electronic program guide that recommends television programs, avoids users getting trapped in a rut when the users select the same programming material over and over again. The recommender may be programmed automatically to leverage the profile of another user to broaden the user's profile. The recommender may use the target descriptions of other users in a same household of the user as a guide for broadening the user's profile. Alternatively, the household profile may be used as a filter for source material for soliciting feedback from the user. In this way, rather than simply broadening the user's range arbitrarily, guidance from other profiles, related in some way to the user, is obtained and leveraged. Note that the "relationship" can include friends, published stereotypes representing interests of the user, and others.

WO 03/030528 A3

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International PCT/IB 02/03693	Location No
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 G06F17/30 H04N7/16			
According to International Patent Classification (IPC) or to both national classification and IPC			
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 G06F H04N			
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched			
Electronic data base consulted during the International search (name of data base and, where practical, search terms used) EPO-Internal, COMPENDEX, INSPEC, PAJ, IBM-TDB, WPI Data, BIOSIS, EMBASE			
C. DOCUMENTS CONSIDERED TO BE RELEVANT			
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	
X	US 6 005 597 A (SIMPSON M O S C JASBA ET AL) 21 December 1999 (1999-12-21) column 5, line 42 -column 8, line 44 column 8, line 45 -column 9, line 10 column 10, line 9 - line 19 figures 7,8	1-10	
Y	---	11	
Y	WO 01 58145 A (KONINKL PHILIPS ELECTRONICS NV) 9 August 2001 (2001-08-09) page 3, line 3 -page 4, line 9 page 7, line 8 -page 8, line 14	11	
E	US 2002/169731 A1 (GUTTA SRINIVAS ET AL) 14 November 2002 (2002-11-14) cited in the application whole document --- -/--	1-11	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.			
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family			
Date of the actual completion of the international search		Date of mailing of the international search report	
25 February 2003		03/03/2003	
Name and mailing address of the ISA European Patent Office, P.O. 5818 Patentean 2 NL - 2200 AD Rijswijk Tel: (+31-70) 340-2040, Tx: 31 651 epo nl Fax: (+31-70) 340-3016		Authorized officer Eichenauer, L	

Form PCT/ISA(210) (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT		International Application No. PCT/IB 02/03693
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	WO 02 17146 A (CABACUNGAN ROBERT SPASSKY ;ENNIS JEFFREY TAYLOR (US); ETANTRUM COM) 28 February 2002 (2002-02-28) page 9, line 23 - line 25 page 12, line 9 - line 18 page 38, line 11 - line 14 page 53, line 6 - line 24 figures 11,19,20 -----	1-10
A	US 5 983 214 A (LANG ANDREW K ET AL) 9 November 1999 (1999-11-09) column 11, line 4 - line 19 column 13, line 5 - line 14 figures 3,4 -----	1-10
A	WO 01 15449 A (SINGULARIS S A ;VAMPARYS FRANCK (CH)) 1 March 2001 (2001-03-01) page 17, line 3 - line 21 figure 7 -----	1-10
A	US 5 974 412 A (NYBAKKEN KRISTOPHER E ET AL) 26 October 1999 (1999-10-26) claim 10 -----	1-10

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.
PCT/IB 02/03693

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6005597	A	21-12-1999	NONE
WO 0158145	A	09-08-2001	WO 0158145 A2 EP 1192802 A2
US 2002169731	A1	14-11-2002	NONE
WO 0217146	A	28-02-2002	AU 2768501 A WO 0217146 A1
US 5983214	A	09-11-1999	US 5867799 A US 2002120609 A1 US 6308175 B1 US 6029161 A US 6314420 B1
WO 0115449	A	01-03-2001	AU 6310500 A WO 0115449 A1
US 5974412	A	26-10-1999	US 6289353 B1

フロントページの続き

(72)発明者 グッタ スリニヴァス ヴィ アール
 オランダ国 5 6 5 6 アー アー アインドーフェン プロフホルストラーン 6
(72)発明者 クラパティ カオシャル
 オランダ国 5 6 5 6 アー アー アインドーフェン プロフホルストラーン 6
F ターム(参考) 5B075 PR03

【要約の続き】

い。