データベースシステム　125－73
（２００１．７．１９）

フィルタリング関数の合成とその性質について

澤井 里枝† 塚本 昌彦† 寺田 努‡ Loh Yin Huei† 西尾 章治郎†

†大阪大学大学院工学研究科情報システム工学専攻
‡大阪大学サイバーメディアセンターサイバーコミュニティ研究部門
{rie,tuka,tsutomu,yhloh,nishio}@ise.eng.osaka-u.ac.jp

近年，さまざまな放送型サービスの普及により，フィルタリング技術に対する要求が高まっている．これまで提案されているフィルタリングシステムでは，ぞれぞれ独自の手法によってデータをフィルタリングしていたが，システムを定性的に表す数学的基盤は存在していなかった．そのため，定性的な評価や最適化，宣言的なフィルタリング言語の設計などができないという問題があった．このような背景から，筆者らはこれまでフィルタリングを関数として表す**フィルタリング関数**を定義し，さまざまなフィルタリングが示す性質間の関係を明らかにしてきた．一般に，実際のフィルタリングシステムにおいて複数のフィルタリング手法を組合せて用いるが，これまでの研究では単一のフィルタリング関数の性質しか明らかにしていなかった．そこで本稿では，フィルタリング関数を合成したときの性質を明らかにすることを目的とする．すなわち，もとの関数の性質が合成関数にどのように影響するかについて，さまざまな組合せを定性的に調べた．これによって合成関数の処理についての指針を示すことができた．

On Composition of Filtering Functions and its Properties

Rie SAWAI† Masahiko TSUKAMOTO† Tsutomu TERADA‡ Yin-Huei LOH† Shojiro NISHIO†

†Department of Information Systems Engineering, Graduate School of Engineering, Osaka University
‡Cybercommunity Division, Cybermedia Center, Osaka University

In recent years, due to the increasing popularization of various broadcast services, there is an increasing demand for filtering techniques. Although filtering systems that have been proposed filter data in their own ways, mathematical representation of these methods does not exist. Consequently, it is not possible to qualitatively evaluate various filtering methods, optimize processing methods in filtering, or design a declarative language for filtering processes. Therefore, in our previous work we have defined *filtering function* that represents filtering as a function, and clarified the relationship between the properties of different filtering. It is common to combine a few methods in current filtering systems in practice, but our previous work revealed only the properties of simple filtering function. In this paper, we clarify the properties of composite function of filtering function. In other words, we found out how the property of original function affects the composite function by qualitatively examining different compositions of functions. Accordingly, we can give a guideline for processing composite function.

1　はじめに

近年，新たな衛星の打ち上げや地上波放送のデジタル化により，多数の放送型サービスが提供されるようになった [7, 9]．このような環境では，放送されるデータ量は膨大になり，内容も広範囲にわたるため，その中からユーザが必要とする情報を探し出すことは非常にコストの高い作業である．そこで，自動的に受信データの取捨選択を行うため，さまざまなフィルタリング機構や，フィルタリングのためのユーザ要求記述言語が提案されている [1, 2, 6, 10, 15]．しかし，各フィルタリング機構は，キーワードマッチングや関連フィードバックなど，ぞれぞれ独自の手法によってデータのフィルタリングを行っているにもかかわらず，それらの手法を定性的に表現する数学的基盤がなかった．そのため，フィルタリングの性質の定性的な評価や処理手法の最適化，宣言的なフィルタリング言語の設計などができなかった．そこで，筆者らはこれまでにフィルタリングを関数として表す**フィルタリング関数**を定義し，フィルタリングが満たす性質をフィルタリング関数が満たす条件によって定性的に表現することを可能にした [11, 13]．さらに，フィルタリング関数の各性質の相互関係を示すことで，さまざまなフィルタリング手法の性質関係を明らかにしている．

筆者らのこれまでの研究では，特定の性質を満たす単一のフィルタリング関数によってフィルタリングを表しているが，実際のフィルタリングシステムは，複数の手法を組合せて実現するのが一般的である．そこで本稿では，複雑なフィルタリングを関数として表現するために，フィルタリング関数の合成について考え，

合成関数の性質を明らかにする．フィルタリング関数の体系に合成の概念を導入することで，複雑な実システムを定性的に表現できる．また，合成関数の相互関係を明らかにすることで，ある性質を満たすシステムが別の性質を満たすシステムと等価であるかどうかが判断でき，環境に応じたより効率的な処理方法への置換が可能となる．

以下，第 2 章でフィルタリング関数の概要を述べ，第 3 章でフィルタリング関数を合成した関数がフィルタリング関数であるための条件について論じる．合成関数の性質関係を第 4 章で示し，特定の処理手法を組合せた場合の性質を第 5 章で明らかにする．そして，第 6 章で証明した合成関数の性質から実際のフィルタリングシステムや関連研究について考察し，第 7 章でまとめを行う．

2　フィルタリング関数

本章では，本研究で基盤とするフィルタリング関数の概要について述べる．まずフィルタリングの処理方法をいくつかのパターンに分類し，フィルタリングを関数の形で表現するフィルタリング関数の定義について述べる．また，分類したフィルタリングの性質をフィルタリング関数が満たす条件として定義する．

2.1　フィルタリング処理の分類

あるフィルタリング手法が与えられたとき，実際の処理方法は以下に示すいくつかのパターンに分類できる．

放送データを受信側にある程度ためておいてから一括してフィルタリングする処理方法を**一括処理**と呼ぶ．それに対し，データアイテムを受信する度に受信データと前回までのフィルタリング結果を合せてフィルタリングする処理方法を**逐次処理**と呼ぶ．また，データ集合を 2 つ以上の任意の集合に分割して各々フィルタリングし，結果をマージしたものをフィルタリング結果とする処理方法を**分配処理**と呼ぶ．さらに，分配処理の結果を再びフィルタリングする処理方法を**並列処理**と呼ぶ．

2.2　フィルタリング関数の性質

データアイテムの集合を $\mathbf{T}$ とする．**フィルタリング関数**とは，任意の $T \subset \mathbf{T}$ に対し[1]，以下の 2 つの条件を満たす $2^{\mathbf{T}}$ 上の関数 f のことをいう．

[1]本稿では $A \subset B$ は A が B の部分集合である (A = B の場合を含む) ことを意味するものとする．

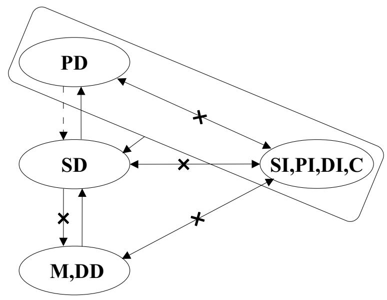

図 1: 性質関係

減少性 (D: Decreasing)

$$f(T) \subset T$$

ベキ等性 (ID: Idempotent)

$$f(f(T)) = f(T)$$

フィルタリング関数について以下のような性質を考える．

単調性 (M: Monotone)

$$S \subset T \text{ ならば } \quad f(S) \subset f(T)$$

分配増加性 (DI: Distributed Increasing)

$$f(S \cup T) \subset f(S) \cup f(T)$$

分配減少性 (DD: Distributed Decreasing)

$$f(S \cup T) \supset f(S) \cup f(T)$$

並列増加性 (PI: Parallel Increasing)

$$f(S \cup T) \subset f(f(S) \cup f(T))$$

並列減少性 (PD: Parallel Decreasing)

$$f(S \cup T) \supset f(f(S) \cup f(T))$$

逐次増加性 (SI: Sequential Increasing)

$$f(S \cup T) \subset f(S \cup f(T))$$

逐次減少性 (SD: Sequential Decreasing)

$$f(S \cup T) \supset f(S \cup f(T))$$

一貫性 (C: Consistency)

$$f(S) \supset f(S \cup T) \cap S$$

ここで，S, T は $\mathbf{T}$ の任意の部分集合とする．

これまでに筆者らはこれらの性質間に，図 1に示すような相互関係があることを明らかにしている．図 1 の矢印は包含関係を表し，一つの枠内に列記した性質は同値であることを示す．さらに，異なる性質を囲った枠はそれらの性質を全て満たす性質を表す．例えば M を満たすフィルタリング関数は SD も満たし，PD かつ SI を満たすフィルタリング関数は SD を満たす．また，M と DD は同値である．ただし，PD を満たすフィルタリング関数が SD を満たすかどうかは現在のところまだ明らかとなっていないため点線で示している．

表 1: ベキ等性を満たさない合成関数

x	$f(x)$	$g(x)$	$f(g(x))$
ϕ	ϕ	ϕ	ϕ
$\{a\}$	$\{a\}$	ϕ	ϕ
$\{b\}$	$\{b\}$	$\{b\}$	$\{b\}$
$\{a,b\}$	$\{a\}$	$\{a,b\}$	$\{a\}$

ここで，分配増加性かつ分配減少性を満たす性質を分配等価性と呼び，分配等価性を満たすフィルタリングは一括処理と分配処理の結果が等価となることを意味する．同様に，並列増加性かつ並列減少性を満たす性質を並列等価性と呼び，並列等価性を満たすフィルタリングは一括処理と並列処理の結果が等価となることを意味する．さらに，逐次増加性かつ逐次減少性を満たす性質を逐次等価性と呼び，逐次等価性を満たすフィルタリングは一括処理と逐次処理の結果が等価となることを意味する．これらの性質を用いると，図 1 の性質関係より，ある性質を満たすシステムが別の性質を満たすシステムと等価であるかどうかが判断でき，環境に応じてより効率的な処理方法に置換できる．

3　フィルタリング関数の合成

一般に，実際のフィルタリングシステムは，複数の手法を組合せて実現する．そこで本研究では前章で述べたフィルタリング関数の合成について考える．本章では，合成した関数がフィルタリング関数であるための条件を示す．

一般に関数 $f : D_1 \to D_2, g : D_2 \to D_3$ に対し，$h(x) = f(g(x))$ で定められる関数 $h : D_1 \to D_3$ をつくることを f と g との合成といい，$h = f \circ g$ と表す．h を f と g の合成関数という．また，$f : D_1 \to D_2$ のとき $Im(f) \triangleq \{f(X)|X \in D_1\}$ を f の値域とよぶ．

フィルタリング関数の合成関数は必ずしもフィルタリング関数にならない．表 1に合成関数がベキ等性を満たさないフィルタリング関数の例を示す．

フィルタリング関数 f, g に対して，f **が** g **にフィルタリング合成可能**であるとは，合成関数 $f \circ g$ がフィルタリング関数であることをいう．一般に，f が g にフィルタリング合成可能であることと g が f にフィルタリング合成可能であることは一致するとはかぎらない．例えば，表 1の f, g で前者は成り立たないが後者は成り立つ．また，f **が** g **不変**であるとは，任意の $X \in Im(f \circ g)$ に対して $f(X) = g(X)$ が成立することをいう．このとき以下の定理が成り立つ．

[定理 1] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であることと f が g 不変であることは同値である．

≪ 証明 ≫

($\Longrightarrow$) f が g にフィルタリング合成可能であるとき，$\exists X_0 \subset \mathbf{T}, Y_0 = f(g(X_0))$ に対して

$$f(Y_0) \neq g(Y_0) \tag{1}$$

と仮定する．このとき，

$$Y_0 \neq f(Y_0), \quad Y_0 \neq g(Y_0) \tag{2}$$

の少なくとも一方を満たす．したがって f, g が減少性を満たすことから，

$$\begin{aligned} Y_0 &\neq f(g(Y_0)) \\ f(g(X_0)) &\neq f(g(f(g(X_0)))) \end{aligned} \tag{3}$$

となり，これは $f \circ g$ のベキ等性を満たさないので，f が g にフィルタリング合成可能であることに反する．

($\Longleftarrow$) まず，$\forall X \in Im(f \circ g)$ に対して $f(X) = g(X)$ ならば

$$X = f(X) = g(X) \tag{4}$$

であることを示す．$\exists X_0 \subset \mathbf{T}, Y_0 = f(g(X_0))$ に対して

$$Y_0 \neq f(Y_0) = g(Y_0) \tag{5}$$

と仮定すると，

$$f(g(X_0)) \neq f(f(g(X_0))) \tag{6}$$

となる．これはベキ等性を満たさないので，f がフィルタリング関数であることに反する．したがって (4) が成立する．

$\forall X \in Im(f \circ g)$ に対して $X = f(X) = g(X)$ ならば $X = f(g(X))$ が成立するので，合成関数 $f \circ g$ はベキ等性を満たす．また，f, g は減少性を満たすので，$X \supset g(X) \supset f(g(X))$ が成立する．したがって，$f \circ g$ は減少性も満たす． □

各性質を満たすフィルタリング関数がフィルタリング合成可能であるかについて以下の定理が成り立つ．

[定理 2] フィルタリング関数 f, g に対して，g が SI(またはそれと等価な PI，DI，C) を満たすならば，f が g にフィルタリング合成可能である．

≪ 証明 ≫ g が C を満たすときを考える．$S \subset \mathbf{T}$ に対して $X = g(S), Y = f(X)$ とすると，

$$\begin{aligned} g(X) = g(g(S)) = g(S) = X \quad (\because \text{ID}) \\ f(Y) = f(f(X)) = f(X) = Y \quad (\because \text{ID}) \end{aligned}$$

となる．S, X, Y は減少性より $S \supset X \supset Y$ を満たすので，

$$g(Y) \supset g(Y \cup S) \cap Y = g(S) \cap Y = X \cap Y = Y$$

が成立する．また，減少性より $g(Y) \subset Y$ なので $g(Y) = Y$ が導かれる．したがって $f(Y) = g(Y)$．f が g 不変であることが示されたので定理 1 より f が g にフィルタリング合成可能である． □

逆が成り立たないことは容易に確かめられる．

フィルタリング関数 f, g に対して，g が SI(または PI，DI，C) を満たさない関数であるとき，f が g にフィルタリング合成可能であるかは，定理 1 より f が g 不変であることを調べればよい．

4 合成関数の性質

定義したフィルタリング関数の性質のうち，図 1に示したようにM，SI，SD，PD の 4 つの性質が同値ではない．本章では，これらの同値ではない性質の全ての組合せについて合成関数を考え，それらの性質を明らかにするため，以下の補題を示す．補題の証明は付録に示す．

[補題 1] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f, g が M を満たすならば，$f \circ g$ は M を満たす． □

[補題 2] フィルタリング関数 f, g に対して，f が M，g が SI を満たすならば，$f \circ g$ で M を満たさないもの，および SI を満たさないものが存在する． □

[補題 3] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f が M，g が SD を満たすならば，$f \circ g$ で M を満たさないもの，および SD を満たさないものが存在する． □

[補題 4] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f が M，g が PD を満たすならば，$f \circ g$ で M を満たさないもの，および PD を満たさないものが存在する． □

[補題 5] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f が SI，g が M を満たすならば，$f \circ g$ で SI を満たさないもの，および M を満たさないものが存在する． □

[補題 6] フィルタリング関数 f, g に対して，f, g が SI を満たすならば，$f \circ g$ で SI を満たさないものが存在する． □

[補題 7] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f が SI，g が SD を満たすならば，$f \circ g$ で SI を満たさないもの，および SD を満たさないものが存在する． □

[補題 8] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f が SI，g が PD を満たすならば，$f \circ g$ で SI を満たさないもの，および PD を満たさないものが存在する． □

[補題 9] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f が SD，g が M を満たすならば，$f \circ g$ で SD を満たさないもの，および M を満たさないものが存在する． □

[補題 10] フィルタリング関数 f, g に対して，f が SD，g が SI を満たすならば，$f \circ g$ で SD を満たさないもの，および SI を満たさないものが存在する． □

[補題 11] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f, g が SD を満たすならば，$f \circ g$ で SD を満たさないものが存在する． □

[補題 12] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f が SD，g が PD を満たすならば，$f \circ g$ で SD を満たさないもの，および PD を満たさないものが存在する． □

[補題 13] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f が PD，g が M を満たすならば，$f \circ g$ で PD を満たさないもの，および M を満たさないものが存在する． □

[補題 14] フィルタリング関数 f, g に対して，f が PD，g が SI を満たすならば，$f \circ g$ で PD を満たさないもの，および SI を満たさないものが存在する． □

[補題 15] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f が PD，g が SD を満たすならば，$f \circ g$ で PD を満たさないもの，および SD を満たさないものが存在する． □

[補題 16] フィルタリング関数 f, g に対して，f が g にフィルタリング合成可能であり，f, g が PD を満たすならば，$f \circ g$ で PD を満たさないものが存在する． □

表 2: 合成関数の性質

f \ g	M	SI	SD	PD
M	M	¬M, ¬SI	¬M, ¬SD	¬M, ¬PD
SI	¬M, ¬SI	¬SI	¬SI, ¬SD	¬SI, ¬PD
SD	¬M, ¬SD	¬SI, ¬SD	¬SD	¬SD, ¬PD
PD	¬M, ¬PD	¬SI, ¬PD	¬SD, ¬PD	¬PD

以上の補題から，全ての性質の組合せによる合成関数の性質をまとめたものを表 2に示す．表 2中の各要素は f, g がそれぞれ行，列の性質をもち，f が g にフィルタリング合成可能であるとき，合成関数 $f \circ g$ が満たす性質を表す．また，¬ はその性質を必ずしも満たさないことを表す．表 2より，単調性を満たすフィルタリング関数どうしを合成したときのみ単調性が保たれることがわかる．一方，それ以外の性質をもつフィルタリング関数の合成では，どちらの性質も保たれない．

5 等価性を満たすフィルタリング関数の合成

前章では，単一の性質のみを満たすフィルタリング関数について論じたが，本章では，分配等価性や逐次等価性，並列等価性を満たすフィルタリング関数を合成した場合の性質を明らかにする．

等価性を満たすフィルタリング関数の合成についての補題を以下に示す．

[補題 17] フィルタリング関数 f, g に対して，f, g が分配等価性を満たすならば，$f \circ g$ は分配等価性を満たす．

≪ 証明 ≫ 分配等価性を満たす f, g に対して

$$f(g(S \cup T)) = f(g(S)) \cup f(g(T)) \tag{7}$$

を示せばよい．文献 [12] より

$$\exists X, \forall S, f(S) = S \cap X \iff \forall S, \forall T, f(S \cup T) = f(S) \cup f(T)$$
$$\exists Y, \forall S, g(S) = S \cap Y \iff \forall S, \forall T, g(S \cup T) = g(S) \cup g(T)$$

が成立するので，$X = f(\mathbf{T}), Y = g(\mathbf{T})$ とおくと，$\forall A \subset \mathbf{T}$ について $f(A) = A \cap X, g(A) = A \cap Y$ とおける．(7) の左右の辺はそれぞれ

$$\begin{aligned} f(g(S \cup T)) &= f((S \cup T) \cap Y) \\ &= f((S \cap Y) \cup (T \cap Y)) \end{aligned}$$

$$
\begin{aligned}
&= ((S\cap Y)\cup(T\cap Y))\cap X \\
&= (S\cap Y\cap X)\cup(T\cap Y\cap X) \quad (8)\\
f(g(S))\cup f(g(T)) &= f(S\cap Y)\cup f(T\cap Y)\\
&= (S\cap Y\cap X)\cup(T\cap Y\cap X) \quad (9)
\end{aligned}
$$

となる．したがって (8),(9) より

$$f(g(S\cup T)) = f(g(S))\cup f(g(T))$$ □

[補題 18] フィルタリング関数 f, g に対して，f が分配等価性，g が逐次等価性または並列等価性を満たすならば，$f\circ g$ で分配等価性を満たさないもの，および逐次等価性，並列等価性を満たさないものが存在する． □

[補題 19] フィルタリング関数 f, g に対して，f が逐次等価性または並列等価性, g が分配等価性を満たすならば，$f\circ g$ は逐次等価性と並列等価性を満たす．

≪ 証明 ≫ 図 1より逐次等価性と並列等価性が同値であるので，逐次等価性を満たす f，分配等価性を満たす g に対して

$$f(g(S\cup T)) = f(g(S\cup f(g(T)))) \quad (10)$$

を示せばよい．文献 [12] より

$$\exists X, \forall S, g(S) = S\cap X \Longleftrightarrow \forall S, \forall T, g(S\cup T) = g(S)\cup g(T)$$

が成立するので，$X = g(\mathbf{T})$ とおくと，$\forall A\subset \mathbf{T}$ について $g(A) = A\cap X$ とおける．(10) の左右の辺はそれぞれ

$$
\begin{aligned}
f(g(S\cup T)) &= f((S\cup T)\cap X)\\
&= f((S\cap X)\cup(T\cap X))\\
&= f((S\cap X)\cup f(T\cap X)) \quad (11)\\
f(g(S\cup f(g(T)))) &= f((S\cup f(T\cap X))\cap X)\\
&= f((S\cap X)\cup(f(T\cap X)\cap X)) \quad (12)
\end{aligned}
$$

となる．ここで次の式が成立することを示す．

$\forall T, \forall X, f(T\cap X) = f(T\cap X)\cap X$

≪ 証明 ≫

i) $f(T\cap X)\supset f(T\cap X)\cap X$ であることは自明．

ii) $f(T\cap X)\subset f(T\cap X)\cap X$ を示す．

$$\forall x\in f(T\cap X)\subset T\cap X\subset X$$

$$\therefore x\in f(T\cap X)\cap X$$

i), ii) より題意は示された． □

ゆえに，(12) は

$$f(g(S\cup f(g(T)))) = f((S\cap X)\cup f(T\cap X))) \quad (13)$$

となる．したがって (11),(13) より

$$f(g(S\cup T)) = f(g(S\cup f(g(T))))$$ □

[補題 20] フィルタリング関数 f, g に対して，f が逐次等価性または並列等価性，g が分配等価性を満たすならば，$f\circ g$ で分配等価性を満たさないものが存在する． □

表 3: 等価性を満たすフィルタリング関数の合成

$f\setminus g$	分配等価性	逐次・並列等価性
分配等価性	分配等価性	–
逐次・並列等価性	逐次・並列等価性	–

[補題 21] フィルタリング関数 f, g に対して，f, g が逐次等価性または並列等価性を満たすならば，$f\circ g$ で逐次等価性を満たさないもの，および並列等価性を満たさないものが存在する． □

以上の補題から，各等価性を満たすフィルタリング関数を合成した場合，合成関数がどの等価性を満たすかをまとめたものを表 3に示す．これより，分配等価性を満たすフィルタリング関数を適用した後は，分配等価性，逐次等価性，並列等価性のいずれの性質を満たすフィルタリング関数を適用した場合でも，後に適用した関数の等価性を保つことが示された．

6 考察

本章では，実際に用いられているいくつかのフィルタリング手法を取り上げ，本稿で述べた合成関数に適用した場合の性質について議論する．

6.1 合成システムの性質

キーワードと論理演算子によってフィルタリングするシステムは，フィルタリングするデータアイテム集合にかかわらず，特定のキーワードを含むデータアイテムは蓄積され，含まなければ蓄積されないので単調性を満たす．また，コンテンツに有効期限を設定し，期限切れしたものから削除するシステムも同様に単調性を満たす．したがって，このようなシステムを合成したシステムは単調性を満たす．一方，単調性を満たさないフィルタリングシステムの場合，どのような組合せで合成しても，本稿で扱った性質はいずれも満たさない．

上記のようなフィルタリングシステムは，データアイテム間の相関性を考慮せず，一貫性も満たすため分配等価性をもつ．このような性質をもつフィルタリングシステムどうしを合成すると，分配等価性を満たすため，一括処理，分配処理，逐次処理，並列処理の結果が全て等価となる．

データのランキングを作成し，重要度の高いものから 10 個のデータを蓄積するフィルタリングシステムの場合，データを分配し，受信機ごとにベスト 10 を選択すると，結果として受信機数分のランキングデータが蓄積されるので分配等価性を満たさない．しかし，データ間に相関性がなく，フィルタリングする集合が大きいほど蓄積されないデータが増加するので一貫性を満たす．また，どの処理段階においても上位 10 個のデータは必ず残され，最後に全体の結果をフィルタリングすれば上位 10 個以外のデータは蓄積されないので逐次等価性を満たす．

上記のような逐次等価性を満たすフィルタリングシステムどうしを合成しても，逐次等価性を満たさない．したがって，分配処理や逐次処理，並列処理の結果と一括処理の結果が等価となることを保証できない．

一方，分配等価性を満たすシステムの後に逐次等価性を満たすシステムを組合せると，逐次等価性を満たすため，一括処理，逐次処理，並列処理の結果が等価となる．逆に，逐次等価性を満たすシステムの後に分配等価性を満たすシステムを組合せると，どちらの等価性も保たれないことから，分配処理や逐次処理，並列処理の結果と一括処理の結果が等価となることを保証できない．

6.2 関連研究への適用

キーワードマッチングによってフィルタリングするシステムに，ニュースグループへ適用する INFOSCOPE[3] や，リアルタイムに音楽を再生する Lyric-Time[5] などがある．これらはプロファイルが更新されない限り単調性や一貫性を満たすため，分配等価性を満たす．距離の公理を用いたフィルタリングシステム [4] は，STB(set top box) によってデータを受信するもので，放送するデータアイテムをその内容に応じて距離空間に置き，ユーザの嗜好を表した点に近いものを蓄積する．これもデータ間の相関性を考慮せず，ユーザの嗜好との距離が一定以下であれば必ず蓄積するため，単調性や一貫性をもち，分配等価性を満たす．したがって，これらのシステムどうしを合成した場合も分配等価性を満たすため，一括処理と分配処理，逐次処理，並列処理のいずれを行っても結果は等価となる．ゆえに，ネットワークの帯域に充分余裕がある場合，一括処理を選択することでサーバの負荷を低減でき，ネットワークが混んできた場合は，複数のサイトから同時にデータを受信する方法に切替えることができる．また，STB のように処理能力が低い受信機を用いて処理する場合は一括処理，比較的記憶容量の小さな受信機の場合は逐次処理が向いている．さらに，複数台の受信機を設置できるのであれば，分配処理や並列処理によって負荷を分散できる．

ProfBuilder[14] は，コンテンツによるフィルタリングをオプションとして選択すると，ユーザのプロファイルに応じたランキングを提示する．このシステムはプロファイルが変更されない間は，一貫性かつ逐次減少性を満たすため，逐次等価性を満たす．したがって，一括処理，逐次処理，並列処理の結果が等価となる．ゆえに，このような性質を示すシステムどうしを組合せた場合，逐次等価性を満たさないため，分配処理や逐次処理，並列処理の結果と一括処理の結果が等価となることを保証できない．つまり，フィルタリング結果の等価性を維持しながら，処理の途中で処理手法を変換することができない．したがって，実装の段階でフィルタリング環境を充分調査し，最適な処理手法を選択しておくことが必要である．しかし，分配等価性を満たすフィルタリングシステムの後に ProfBuilder を組合せると，逐次等価性や並列等価性を満たすため，一括処理，逐次処理，並列処理の結果が等価となる．ゆえに，逐次処理をすることで常にランキングを最新の状態にしておける．また，並列処理をすることでネットワークの使用帯域や受信機の処理コストを小さくできる．

放送データをフィルタリングするシステムに AIS (Active Information Store) [8] がある．これは，まず受信したデータをキーワードマッチングによって絞り込み，次に評価値を計算してディスク容量の制限に合せて蓄積するデータを決定する．前者の処理は単調性と一貫性を満たすので分配等価性を満たし，後者の処理は一貫性かつ逐次減少性を満たすので逐次等価性を満たす．したがって，AIS は分配等価性を満たすフィルタリングの後に逐次等価性を満たすフィルタリングを組合せたシステムであり，システム全体は逐次等価性や並列等価性を満たす．ゆえに，一括処理，逐次処理，並列処理の結果が等価となる．したがって，データがある程度たまってから処理することで受信機の負荷を軽減できる．また，多数の放送チャンネルが存在する場合は 2 つの受信機で並列処理が可能である．

7 おわりに

本稿では，フィルタリング関数の合成関数がフィルタリング関数であるための条件を示し，さまざまな性質を満たすフィルタリング関数を合成したときの性質を明らかにした．フィルタリング関数の体系に合成関

数を導入したことで，複雑なフィルタリングシステムを定性的に表現することが可能となった．また，実際のフィルタリングシステムをその性質によって分類し，等価置換が可能な処理方法について述べた．本稿で示した体系を実システムに適用することで，環境に応じてより効率的なフィルタリング処理を選択できる．

今後の課題は以下のようなものが挙げられる．

- 合成関数への条件の付加
  本稿で論じたほとんどの合成関数は必ずしももとの関数の性質が保たれないことから，実際のフィルタリングシステムを組合せる場合に注意が必要である．しかし，フィルタリング関数を合成する際，特定の条件を追加し，もとのフィルタリング関数が満たす性質の範囲を狭くすることで合成関数の性質が保たれる可能性がある．
- 特定の処理を表す関数の定義
  キーワードによる抽出やランキング作成など，実際のフィルタリングシステムで広く行われている手法については，それのみを表現する特別な関数を定義する．関数が満たす性質の範囲を狭めることで，特定の処理に焦点を絞った議論ができる．

謝辞

本研究は，日本学術振興会基盤研究(B) (2) (12480095)および日本学術振興会未来開拓学術研究推進事業における研究プロジェクト「マルチメディア・コンテンツの高次処理の研究 (Project No. JSPS-RFTF97P00501) の研究助成によるものである．ここに記して謝意を表す．

参考文献

[1] N. J. Belkin and W. B. Croft: “Information filtering and information retrieval: two sides of the same coin?,” *Communications of the ACM*, vol. 35, no. 12, pp. 29–38 (1992).

[2] T. A. H. Bell and A. Moffat: “The design of a high performance information filtering system,” in *Proc. SIGIR '96*, pp. 12–20 (1996).

[3] G. Fischer and C. Stevens: “Information access in complex, poorly structured information spaces,” in *Proc. Human Factors in Computing Systems CHI'91 Conference*, pp. 63–70 (1991).

[4] カンギョウビ, 大和田俊和, 浅田一繁, 飯沢篤志, 古瀬一隆: “情報放送システムにおける距離の近似を利用したフィルタリング方式,” 電子情報通信学会第 11 回データ工学ワークショップ (DEWS2000) 論文集 (CD-ROM) (2000).

[5] S. Loeb: “Architecting personalized delivery of multimedia information,” *Communications of the ACM*, vol. 35, no. 12, pp. 39–48 (1992).

[6] 森田昌宏: “情報フィルタリングに関する研究動向,” JAIST Research Report, IS-RR-93-9I, 北陸先端科学技術大学院大学情報科学研究科 (1993).

[7] 西正, 野村敦子: “多チャンネル放送の衝撃,” 中央経済社 (1997).

[8] S. Sanguantrakul, T. Terada, M. Tsukamoto, S. Nishio, K. Miura, S. Matsuura, and T. Imanaka: “A user customized selection and categorization for broadcast data,” in *Proc. 1999 IEEE International Workshops on Multimedia Network Systems(MMNS)*, pp. 596–601 (1999).

[9] サテライトマガジン: http://www.satemaga.co.jp/.

[10] 澤井里枝, 寺田努, 塚本昌彦, 西尾章治郎: “フィルタリング SQL: フィルタリングのためのユーザ要求記述言語,” 電子情報通信学会第 11 回データ工学ワークショップ (DEWS2000) 論文集 (CD-ROM) (2000).

[11] 澤井里枝, Loh Yin Huei, 寺田努, 塚本昌彦, 西尾章治郎: “フィルタリングの関数的性質とその関係について,” 電子情報通信学会第 12 回データ工学ワークショップ (DEWS2001) 論文集 (CD-ROM) (2001).

[12] 澤井里枝, 塚本昌彦, 寺田努, Loh Yin Huei, 西尾章治郎: “Selection 関数について,” 西尾研ワーキングメモ (2001).

[13] R. Sawai, M. Tsukamoto, Y. H. Loh, T. Terada, and S. Nishio: “Functional properties of information filtering,” in *Proc. 27th International Conference on Very Large Data Bases (VLDB2001)*, to appear (2001).

[14] A. M. A. Wasfi: “Collecting user access patterns for building user profiles and collaborative filtering,” in *Proc. 1999 International Conference on Intelligent User Interfaces*, pp. 57–64 (1999).

[15] T. W. Yan and H. Garcia-Molina: “The SIFT information dissemination system,” *ACM Transactions on Database Systems*, vol. 24, no. 4, pp. 529–565 (1999).

付録

≪ 補題 1 の証明 ≫ $\forall S, \forall T \subset \mathbf{T}, S \supset T$ ならば $f(S) \supset f(T), g(S) \supset g(T)$. ゆえに $f(g(S)) \supset f(g(T))$. □

≪ 補題 2 の証明 ≫ $\mathbf{T} = \{a, b\}$ とする．表4に示すフィルタリング関数は $\forall S, \forall T \subset \mathbf{T}$ に対して，f は M，g は SI を満たすが，$S = \{a, b\}, T = \{b\}$ のとき $f(g(S)) \supset f(g(T))$ を満たさない．

また，表5に示すフィルタリング関数は $\forall S, \forall T \subset \mathbf{T}$ に対して，f は M，g は SI を満たすが，$S = \{b\}, T = \{a\}$ のとき $f(g(S \cup T)) \subset f(g(S \cup f(g(T))))$ を満たさない． □

≪ 補題 3 の証明 ≫ $\mathbf{T} = \{a, b\}$ とする．表4に示すフィルタリング関数は $\forall S, \forall T \subset \mathbf{T}$ に対して，f は M，g は SD を満たすが，$S = \{a, b\}, T = \{b\}$ のとき $f(g(S)) \supset f(g(T))$ を満たさない．

また，$S = \{b\}, T = \{a\}$ のとき $f(g(S \cup T)) \supset f(g(S \cup f(g(T))))$ を満たさない． □

≪ 補題 4 の証明 ≫ 省略 (補題 3 と同様に証明できる)． □

表 4: 反例 1

x	$f(x)$	$g(x)$	$f(g(x))$
ϕ	ϕ	ϕ	ϕ
$\{a\}$	ϕ	$\{a\}$	ϕ
$\{b\}$	$\{b\}$	$\{b\}$	$\{b\}$
$\{a,b\}$	$\{b\}$	$\{a\}$	ϕ

表 5: 反例 2

x	$f(x)$	$g(x)$	$f(g(x))$
ϕ	ϕ	ϕ	ϕ
$\{a\}$	ϕ	$\{a\}$	ϕ
$\{b\}$	$\{b\}$	$\{b\}$	$\{b\}$
$\{a,b\}$	$\{a,b\}$	$\{a,b\}$	$\{a,b\}$

表 6: 反例 3

x	$f(x)$	$g(x)$	$f(g(x))$
ϕ	ϕ	ϕ	ϕ
$\{a\}$	$\{a\}$	$\{a\}$	$\{a\}$
$\{b\}$	$\{b\}$	$\{b\}$	$\{b\}$
$\{c\}$	$\{c\}$	$\{c\}$	$\{c\}$
$\{a,b\}$	$\{a,b\}$	ϕ	ϕ
$\{a,c\}$	$\{a\}$	$\{a,c\}$	$\{a\}$
$\{b,c\}$	$\{b,c\}$	$\{b,c\}$	$\{b,c\}$
$\{a,b,c\}$	$\{a,b\}$	$\{c\}$	$\{c\}$

表 7: 反例 4

x	$f(x)$	$g(x)$	$f(g(x))$
ϕ	ϕ	ϕ	ϕ
$\{a\}$	$\{a\}$	ϕ	ϕ
$\{b\}$	$\{b\}$	$\{b\}$	$\{b\}$
$\{a,b\}$	$\{a,b\}$	$\{a,b\}$	$\{a,b\}$

表 8: 反例 5

x	$f(x)$	$g(x)$	$f(g(x))$
ϕ	ϕ	ϕ	ϕ
$\{a\}$	$\{a\}$	$\{a\}$	$\{a\}$
$\{b\}$	$\{b\}$	$\{b\}$	$\{b\}$
$\{a,b\}$	$\{a\}$	$\{a,b\}$	$\{a\}$

表 9: 反例 6

x	$f(x)$	$g(x)$	$f(g(x))$
ϕ	ϕ	ϕ	ϕ
$\{a\}$	$\{a\}$	$\{a\}$	$\{a\}$
$\{b\}$	$\{b\}$	ϕ	ϕ
$\{c\}$	$\{c\}$	$\{c\}$	$\{c\}$
$\{a,b\}$	$\{a,b\}$	$\{a\}$	$\{a\}$
$\{a,c\}$	$\{a,c\}$	$\{a,c\}$	$\{a,c\}$
$\{b,c\}$	$\{b,c\}$	$\{b,c\}$	$\{b,c\}$
$\{a,b,c\}$	$\{b,c\}$	$\{a,b,c\}$	$\{b,c\}$

$\ll$ **補題 5 の証明** $\gg$ $\mathbf{T}=\{a,b\}$ とする．表 7に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，f は SI，g は M を満たすが，$S=\{b\},T=\{a\}$ のとき，$f(g(S\cup T))\subset f(g(S\cup f(g(T))))$ を満たさない．

また，表 8に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，f は SI，g は M を満たすが，$S=\{a,b\},T=\{b\}$ のとき $f(g(S))\supset f(g(T))$ を満たさない．□

$\ll$ **補題 6 の証明** $\gg$ $\mathbf{T}=\{a,b,c\}$ とする．表 6に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，f,g は SI を満たすが，$S=\{a,b\},T=\{a,c\}$ のとき $f(g(S\cup T))\subset f(g(S\cup f(g(T))))$ を満たさない．□

$\ll$ **補題 7 の証明** $\gg$ $\mathbf{T}=\{a,b\}$ とする．表 7に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，f は SI，g は SD を満たすが，$S=\{b\},T=\{a\}$ のとき，$f(g(S\cup T))\subset f(g(S\cup f(g(T))))$ を満たさない．

また，表 4に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，f は SI，g は SD を満たすが，$S=\{b\},T=\{a\}$ のとき，$f(g(S\cup T))\supset f(g(S\cup f(g(T))))$ を満たさない．□

$\ll$ **補題 8 の証明** $\gg$ 省略 (補題 7 と同様に証明できる)．□

$\ll$ **補題 9 の証明** $\gg$ $\mathbf{T}=\{a,b,c\}$ とする．表 9に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，f は SD，g は M を満たすが，$S=\{a,c\},T=\{a,b\}$ のとき，$f(g(S\cup T))\supset f(g(S\cup f(g(T))))$ を満たさない．

また，$S=\{a,b,c\},T=\{a,b\}$ のとき $f(g(S))\supset f(g(T))$ を満たさない．□

$\ll$ **補題 10 の証明** $\gg$ $\mathbf{T}=\{a,b\}$ とする．表 4に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，f は SD，g は SI を満たすが，$S=\{b\},T=\{a\}$ のとき，$f(g(S\cup T))\supset f(g(S\cup f(g(T))))$ を満たさない．

また，表 5に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，f は SD，g は SI を満たすが，$S=\{b\},T=\{a\}$ のとき，$f(g(S\cup T))\subset f(g(S\cup f(g(T))))$ を満たさない．□

$\ll$ **補題 11 の証明** $\gg$ $\mathbf{T}=\{a,b\}$ とする．表 4に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，f,g は SD を満たすが，$S=\{b\},T=\{a\}$ のとき，$f(g(S\cup T))\supset f(g(S\cup f(g(T))))$ を満たさない．□

$\ll$ **補題 12 の証明** $\gg$ 省略 (補題 11 と同様に証明できる)．□

$\ll$ **補題 13 の証明** $\gg$ 省略 (補題 9 と同様に証明できる)．□

$\ll$ **補題 14 の証明** $\gg$ 省略 (補題 10 と同様に証明できる)．□

$\ll$ **補題 15 の証明** $\gg$ 省略 (補題 11 と同様に証明できる)．□

$\ll$ **補題 16 の証明** $\gg$ 省略 (補題 11 と同様に証明できる)．□

$\ll$ **補題 18 の証明** $\gg$ $\mathbf{T}=\{a,b\}$ とする．表 4に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，$f(x)$ は DI，DD を満たし，$g(x)$ は SI，SD，PI，PD を満たす．これに対し，$S=\{a,b\},T=\{b\}$ のとき，$f(g(S))\supset f(g(T))$，$S=\{b\},T=\{a\}$ のとき，$f(g(S\cup T))\supset f(g(S\cup f(g(T))))$，$f(g(S\cup T))\supset f(g(f(g(S))\cup f(g(T))))$ を満たさない．□

$\ll$ **補題 20 の証明** $\gg$ $\mathbf{T}=\{a,b\}$ とする．表 8に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，$f(x)$ は SI，SD，PI，PD を満たし，$g(x)$ は DI，DD を満たす．これに対し，$S=\{a,b\},T=\{b\}$ のとき，$f(g(S\cup T))\supset f(g(S))\cup f(g(T))$ を満たさない．□

$\ll$ **補題 21 の証明** $\gg$ $\mathbf{T}=\{a,b\}$ とする．表 4に示すフィルタリング関数は $\forall S,\forall T\subset\mathbf{T}$ に対して，$f(x),g(x)$ は SI，SD，PI，PD を満たす．これに対し，$S=\{b\},T=\{a\}$ のとき，$f(g(S\cup T))\supset f(g(S\cup f(g(T))))$，$f(g(S\cup T))\supset f(g(f(g(S))\cup f(g(T))))$ を満たさない．□