東方學報 京都第94冊（2019）：282-268頁

漢字構造の代數的記述について

―― 人文學における形式的思考の一側面 ――

白 須 裕 之

1 は じ め に

これまでの歴史において科學と技術は相補的な役割を演じてきたが，前世紀の終りから情報技術（本稿では計算機，コンピュータなどを使用して，計算を扱うだけでなく，様々な知識の電子的な表現を扱うものも含めて情報技術と呼ぶことにする）は，科學と技術の關係を大きく變えようとしている。現代の學問研究において情報技術の必要性は自明のことに成りつつある。しかし，人文學の場合，情報技術を應用することによる利便性だけに目を奪われていて，學問と技術の相補性に目が向けられていない。特に人文學の對象が既知のものとして扱われ，それを既存の情報技術に押し込めようとするものばかりであり，學問と技術の關係を改めて考察しようという試みは見られない。この狀況はデジタルアーカイブ然り，歴史情報，言語コーパスなどのデータベースやツール，畫像データベース然り，と數え上げればきりがない。

人文學は本來豫め與えられた，既知のものを對象としている譯ではなく，研究の進展とともに新たな對象を創出してきたのは明らかである。從って，これからの人文學と情報技術の相補的な關係を對象という面から考察してみよう。さて情報技術は形式的體系によって對象を操作的に扱うことができることを，その本質的な點とする。ということは，人文學の對象をそれを記述する形式的體系との關係として，再定式化できることを意味する。

形式的體系を學的な體驗の基礎に位置づけ，その定性的な性質を解明したものにG.-G. Granger（1920-2016）の形式的思考がある[1)]。Grangerは「操作」と「對象」の相關關係

1）本稿ではGrangerの學問論が本質的に形式的體系を前提している點を重視して，その內容を「形式的思考」と名付けることにする。本稿に關連するGrangerの思想について，詳し↗

を學的認識の根本に据える。操作の體系と對象の體系は不可分であり，相互規定の關係にある。この關係を「操作と對象の雙對性」と呼ぶ。對象は操作體系の外部からその內容を與えられるのではなく，操作體系それのみが對象に內容を與えることが出來る。卽ち，「操作と對象の雙對性」が「意味一般の根本的な可能性の條件」になる。操作體系の形式化が行なわれるとき，當初の操作體系からはみ出す新たな內容がその形式化から與えられるとき，これを Granger は「形式的內容」と呼ぶ。「形式的內容」は研究の新たな對象として，次の理論の可能性を開く。形式的思考はこのように學知の動的で自律的な發展を支える。

人文學への形式的思考を考える前提として，Granger の思想を誤解を恐れずに單純化して述べてみよう。言語記號として述語論理によるものを使って對象を記述することを考える[2]。對象 a が性質 P を充たしていることを $P(a)$，また二つの對象 a, b が性質 Q を充たしていることを $Q(a, b)$ などと書くことにする。卽ち，a, b や P, Q は注目している學問の領域對象である。もし對象 a, b，述語 P, Q を形式化でき，構文論的な體系が構成できたとすると，a, b や P, Q は形式的體系を充たすものという意味のみを殘して，その痕跡を消し去ることができる。また，その形式的體系を充たす領域對象は a, b や P, Q だけでなく，他の潛在的な領域對象も想定することができるようになる。これは「形式的體系の意味論に關する一般化」と解釋できる[3]。その形式的體系は新たな內容を獲得し，新しい考察の對象へと我々を導くことになる。

本稿ではこのような形式的思考による對象の研究の一側面として，漢字構造を取り上げる。漢字という對象は興味深いことに，人文學の樣々な對象を記述するための言語と關連し，それを書寫するためのある種の單位として出現する。漢字という對象を記述するためには形音義に關わる總合的な考究が必要であり，それについては今後の課題に委ねよう。ここでは形の面を題材として，漢字構造の代數的記述を試み，また，その問題點，課題を提出しようと思う。

漢字構造の代數的記述として，等號系を利用し，その情報の檢證手段として項書換系

↘ くは文獻［5］［6］［9］などを參照されたい。

2） ここでは J. Cavaillès（1903-1944）の「パラディグム」及び「主題化」という概念を Granger が形式的體系との關連でどのように解釋したかを手掛かりとする。ここでの Granger による解釋は文獻［9］を參考にした。形式的體系の構築には樣々な數學的言語を使用することができるが，本稿で利用する等式論理を念頭に，述語論理の例で考察する。等式論理については，等式を使って記述された對象の性質を述語として解釋できる。

3） これが Cavaillès が言う「パラディグム」の Granger による解釋である。また，「主題化」は「形式的體系の構文論に關する一般化」として解釋される。「主題化」の人文學での役割については「おわりに」で觸れる。

を使用する。主に Ideographic Description Sequence（IDS）による漢字構造の記述と包攝關係を取り扱う。漢字構造の形式的な扱いについて，文獻［15］では包攝規準を書換規則として捉え，計算機構として良い性質をもつ項書換系を完備化手續きによって構成するという試みを述べている。しかし，文獻［15］での項書換系の扱いは多くの誤解を含んでいて，折角の試みも充分には展開されていない。

2　基本的な用語

本稿では形式的體系である等號系，項書換系，及び項書換系の完備化手續きを扱う。本節では，そのために必要な等號系と項書換系についての數學的な用語を準備をしよう[4)]。

2.1　代數と等號系

最初に行なうことは本稿での議論に使える記號をきちんと決めておくことである。このような準備は一見迂遠なようでいて，何について議論しているのかを正確に把握するために必要な作業である。

定義 1（言語 language）　言語（或いはシグニチャ *signature*, 類型 *similarity type*）とは $\sigma=\langle\mathcal{F},\rho\rangle$ である。ここで $\mathcal{F}$ は關數記號の集合，ρ は $\mathcal{F}$ から自然數の集合 ω への寫像である。關數記號 $f\in\mathcal{F}$ に對して，自然數 $\rho(f)$ を f のアリティ（またはランク）という。$c\in\mathcal{F}$ に對して，$\rho(c)=0$ のとき c を特に定數記號と呼ぶ。また，$\sigma:\mathcal{F}\to\omega$ のことを言語とも呼ぶ。自然數 n に對して $\mathcal{F}_n$ は集合 $\{f\in\mathcal{F}\mid\rho(f)=n\}$ を意味する。從って $\mathcal{F}_0$ は定數記號の集合である。

「言語」は上で付記したようにその概念を使用する分野に依存して，幾つかの呼び方がある。ここでは言語を省略して，「σ 代數」のように使うこともある。また，以下の議論で代數，項，等號系，項書換系などの概念を使うにあたり，明示しない限り言語は固定して考えるものとする。

定義 2（代數 algebra）　言語 $\sigma=\langle\mathcal{F},\rho\rangle$ の代數（或いは σ 代數）$\mathfrak{A}=\langle A,F\rangle$ とは，空でない

4）　本稿の議論には普遍代數のごく初步の用語を必要とする。普遍代數の基本文獻には［3］［7］などがある。また，計算機科學では項書換系についての理論，應用の研究は非常に盛んであるが，これについては文獻［2］などを參照されたい。項書換系のチュートリアルとしては文獻［10］［14］に目を通すことをお薦めする。

集合 A と關數の集合 F の順序對である。但し，集合 F は關數記號 $f \in \mathcal{F}$ に對應する A 上の關數 $f^{\mathfrak{A}}: A^{\rho(f)} \to A$ からなる集合である。このとき A は $\mathfrak{A}$ の臺集合 *underlying set*（または領域 *domain*），F は基本演算 *fundamental operations* という。

定義 3（準同形寫像 homomorphism）　二つの代數 $\mathfrak{A}, \mathfrak{B}$ に對して寫像 $f: A \to B$ が，任意の $f \in \mathcal{F}_n$ 及び任意の $a_1, \ldots a_n \in A$ に對して $\phi f^{\mathfrak{A}}(a_1, \ldots, a_n) = \phi f^{\mathfrak{B}}$（$\phi a_1, \ldots, \phi a_n$）を充たすとき，$\mathfrak{A}$ から $\mathfrak{B}$ へ準同形寫像といい $\phi: \mathfrak{A} \to \mathfrak{B}$ で示す。

定義 4（項 term）　X を（通常，加算無限個の）變數の集合とする。言語 $\mathcal{F}$ の X 上の項集合 $T(\mathcal{F}, X)$ は以下を充たす：

- 變數 $x \in X$ は項である，
- 任意の $n \geq 0$ について $p_1, \ldots, p_n$ が項であり $f \in \mathcal{F}_n$ であるとき，$f(p_1, \ldots, p_n)$ は項である[5]。

定義 5（文脈 context）　缺落の記號を Ω で示す。文脈は項の定義において，缺落を變數と同樣に使用することで定義される：

- 變數 $x \in X$ または缺落 Ω は文脈である，
- 任意の $n \geq 0$ について $r_1, \ldots, r_n$ が文脈であり $f \in \mathcal{F}_n$ であるとき，$f(r_1, \ldots, r_n)$ は文脈である。

缺落 Ω をもつ文脈 $r[\Omega]$ に對して，その缺落 Ω を項 p で置換えた項を $r[p]$ と書く。同樣にして複數の缺落 $\Omega_1, \ldots, \Omega_m$ をもつ文脈も考えることができる。

定義 6（項代數 term algebra）　$\mathcal{F}$ と X が與えられたとき X 上の項代數 $\mathcal{T}(\mathcal{F}, X)$ は臺集合が項集合 $T(\mathcal{F}, X)$ であり，$f \in \mathcal{F}_n (n \geq 0)$ に對して基本演算 $f^{\mathcal{T}(\mathcal{F},X)}$ は $f^{\mathcal{T}(\mathcal{F},X)}: \langle p_1, \ldots, p_n \rangle \mapsto f(p_1, \ldots, p_n)$ で與えられる。但し，$p_i \in T(\mathcal{F}, X)(1 \leq i \leq n)$ とする。

定義 7（代入 substitution）　變數の集合 X から項集合 $T(\mathcal{F}, X)$ への寫像を代入といい，θ, ϕ などで記す。代入 θ は項代數 $\mathcal{T}(\mathcal{F}, X)$ 間の準同形寫像 $\bar{\theta}$ に自然に擴張できる。

定義 8（割當 assignment）　代數 $\mathfrak{A} = \langle A, F \rangle$ が與えられたとき，變數の集合 X から臺集合 A

5）　$n=0$ のとき定數 $c \in \mathcal{F}_0$ が項であることに注意。

への寫像 h を割當という。割當は自然に項代數 $\mathcal{T}(\mathcal{F}, X)$ から代數 $\mathfrak{A}$ への準同形寫像 $\bar{h}$ に擴張できる。項 $t(x_1, \ldots x_n)$ について $\bar{h}$ の値を考えるとき，その値は項 t に含まれる變數 $x_1, \ldots x_n$にしか影響されないので，$h(x_i)=a_i$として，$h(t(x_1, \ldots x_n))$を $t^{\mathfrak{A}}(a_1, \ldots a_n)$ と記す。

定義 9（等號系 equational system）　言語 $\mathcal{F}$ の二つの項 p, q を等號 $\approx$ で結んだ $p\approx q$ を等式とよぶ[6]。等式の有限集合 E を等號系と呼ぶ。

定義 10（モデル model）　等式 $p\approx q$ が代數 $\mathfrak{A}$ で恒眞 *valid* であるとは，任意の割當 θ に對して $\bar{\theta}(p)=\bar{\theta}(q)$ が成り立つことをいう。これを $\mathfrak{A} \vDash p\approx q$ で示す。また等號系 E の任意の等式が代數 $\mathfrak{A}$ で恒眞であるとき，$\mathfrak{A}$ は等號系 E のモデルであるという。

ある等式が E に含まれる等式から等式論理の推論によって導かれるならば，その等式は E のモデルでも恒眞である。逆にモデルとなる代數の性質を記述する上で等號系は有效な方法になっている[7]。等式論理の推論によって導かれる等式を機械的に檢證する方法の一つに項書換系による方法があるが，以下では項書換系についての事項を確認する。

2.2　簡約系

計算機構の直感的な性質のある側面を捉える爲に，簡約系 reduction system についてまず纏めておく。

簡約系 $\langle A, \rightarrow\rangle$ はある集合 A とその上の關係→からなる。$a \rightarrow b$ であるならば何らかの意味で，a と b が等しく，b は a より簡單になっているという直感的な意味を表現する。關係→の反射推移閉包 reflexive transitive closure を$\xrightarrow{*}$と書く。このとき $\langle \mathrm{A}, \xrightarrow{*}\rangle$ は半順序集合 preordered set になる，卽ち以下を充たす：

・反射律：$a \xrightarrow{*} a,$

・推移律：$a \xrightarrow{*} b, b \xrightarrow{*} c$ ならば $a \xrightarrow{*} c.$

また，關係→の反射推移對稱閉包 reflexive transitive symmetric closure を $\sim$ で示す。これは簡約關係→を媒介として表現される等しいという關係を示す。

簡約系を計算機構として見た場合に重要な性質が停止性と合流性である。停止性は

6） ここで等式に記號≈を使うのは，$p\approx q$ が統語論的な意味での記號であり，數學的な意味の等しいことを示す記號と區別するためである。記號はしばしば近似の意味に使用されるが，ここではそのような意味はない。

7） 等式論理と代數との關連について，詳しくは普遍代數の文獻［3］［7］などを參照されたい。

$a_0 \to a_1 \to a_2 \to \cdots$のような無限の簡約列が存在しないこと，即ち簡約という計算が無限に續かないことを意味する。$a \to b$ なる b が存在するとき a を可約 reducible といい，そうでないとき既約 irreducible であるという。任意の元 a から簡約を行なって，既約元に達するときそれを a の既約形 irreducible form といい $a\downarrow$ で示す[8]。

停止性を以下の整礎集合という概念を使用して定義すると，$\langle A, \overset{*}{\to}\rangle$ が停止性をもつことと，整礎集合であるとは同値であることが言える。

定義 11（整礎集合 well-founded set） 順序集合 $\langle A, \geq\rangle$ が整礎集合であるとは，A の空でない任意の部分集合が極小元をもつことをいう。$a \in A$ が極小元であるとは，$a > b$ なる元 b が存在しないことをいう。

合流性とは $a \overset{*}{\to} b_1, a \overset{*}{\to} b_2$ のとき $b_1 \overset{*}{\to} c, b_2 \overset{*}{\to} c$ なる元 c が存在することをいう。また，Church–Rosser 性とは $a \sim b$ のとき $a \overset{*}{\to} c, b \overset{*}{\to} c$ なる元 c が存在することをいう。合流性と Church–Rosser 性は同値の概念である。

2.3 項書換系

定義 12（項書換系 term rewriting system） 項集合 $T(\mathcal{F}, X)$ 上に簡約$\to$が定義されているとき，$\langle T(\mathcal{F}, X), \to\rangle$ を項書換系という。多くの場合，簡約$\to$は書換規則の有限集合 R から以下のように導出される。項 $p, q \in T(\mathcal{F}, X)$ に對して文脈 $r[\Omega]$，代入 θ，書換規則 $s \Rightarrow t \in R$ が存在して $p = r[\theta(s)], q = r[\theta(t)]$ であるとき，$p \to q$ であるとする。書換規則の集合 R が與えられたとき R を項書換系ということもある。

定義 13（安定 *stable*） 項集合 $T(\mathcal{F}, X)$ 上の關係 ρ が安定であるとは，以下の二つの條件が成り立つことをいう[9]：

1. 任意の文脈 $r[\Omega]$ 及び任意の項 p, q に對して $p\rho q$ ならば $r[p]\rho r[q]$ が成り立つ（文脈に關して兩立する *compatible with context*），

8） 既約形を normal form（正規形）と呼ぶこともあるが，その形に價値判斷を含めない意味でも，ここでは既約形を使用する。

9） 條件 $p\rho q$ ならば $f(p_1, \ldots, p_{i-1}, p, p_{i+1}, \ldots, p_n)\rho f(q_1, \ldots, q_{i-1}, q, q_{i+1}, \ldots, q_n)$（ここで $f \in \mathcal{F}_n$, $p_1, \ldots, p_{i-1}, p, p_{i+1}, \ldots, p_n, q_1, \ldots, q_{i-1}, q, q_{i+1}, \ldots, q_n \in T(\mathcal{F}, X)$ である）が成り立つことを演算に關して閉じている（compatible with operations）という。半順序に對して安定性の條件 1 をこの條件に換えたものを文獻［2］では rewrite order というが，反射律，推移律がある場合にはこれらは同値である。

2. 任意の代入 θ 及び任意の項 p, q に對して $p\rho q$ ならば $\theta(p)\rho\theta(q)$ が成り立つ（代入に關して閉じている *closed under substitution*）.

定義 14（辭書式部分項順序 lexicographic subterm order［10］） 關數記號の集合 $\mathcal{F}$ 上の順序 $>$ に對して以下のように定義される $T(\mathcal{F}, X)$ 上の順序 $\gg$ を辭書式部分項順序と呼ぶ。

(1) 項 $p \in T(\mathcal{F}, X)$ と變數 $x \in X$ に對して $x \gg p$ ではない,

(2) $p=f(p_1, \ldots, p_n)$ に對して $p \gg q$ となるのは以下のいずれかの場合であり，それに限る,

(2. 1) ある i に對して $p_i \gg q$ または $p_i = q$,

(2. 2) $q=g(q_1, \ldots q_m)$ に對して全ての j で $p \gg q_j$ であり，かつ以下のいずれかが成り立つ,

(2. 2. 1) $f > g$,

(2. 2. 2) $f = g$ かつある i について $p_i \gg q_i$ であり，全ての $k<i$ について $p_k=q_k$.

このとき，$\langle T(\mathcal{F}, X), \gg \rangle$ は安定な整礎集合になる。

2. 4 完備化手續き

以下では文獻［2］に基づいて，最も單純な完備化手續きの概要を述べる。ここで述べた手續きは效率には配慮していない。完備化手續きにどのような概念が必要であるかについて，確認することがここでの目的である。その效率化及び適用範圍の擴張など樣々な提案が爲されているが，これについては文獻［2］等を參照されたい。

定義 15（單化子 unifier） 二つの項 p, q に對して $\theta(p)=\theta(q)$ となる代入 θ を p と q の單化子と呼ぶ。代入 θ_1, θ_2 に對して，代入 ϕ が存在して $\theta_1=\phi\theta_2$ のとき $\theta_1 \leq \theta_2$ と書く。項 p, q に對する單化子で，この順序に關して最大の單化子を最汎單化子という。

定義 16（重像 superposition） 變數でない q を部分項としてもつ項 $p[q]$ に對して項 q と r が單化可能であるとき，その最汎單化子を θ とする。このとき $\theta(p[q])$ を $p[q]$ と r の重像という。

定義 17（要對 critical pair） 二つの書換規則 $p[q] \Rightarrow s$ と $r \Rightarrow t$ に對して左邊 $p[q]$ と r が重像 $\theta(p[q])$ をもつとき，對 $\langle \theta(s), \theta(p[t]) \rangle$ を要對という[10]。項書換系 R が與えられた

10） 現在，多くの項書換系の文獻では critical pair の譯語を「危險對」とするが，本稿では文獻↗

とき，その要對の集合を $\mathrm{CP}(R)$ と書く。

重像は書換規則と $p[q]$ の部分項 q の選び方に依存する。重像 $\theta(p[q])$ は二つの書換規則で各々，對 $\theta(s)$ と $\theta(p[t])$ の二つに書換られる。全ての要對が項書換系で同一の項に簡約できることが R が合流性をもつことの必要十分條件である。從って，以下のような完備化手續きを構成できる。

定義 18（完備化手續き［2]） 項集合 $T(\mathcal{F}, X)$ とその上の安定な整礎順序 $\leq$ 及び等號系が與えられたとき，完備な項書換系を求める完備化手續きは以下である[11]。

入　力：$T(\mathcal{F}, X)$ 上の等號系 E 及び簡約順序 $>$。

出　力：手續きが成功裏に停止したとき E と等價な有限で完備な項書換系 R，そうでない場合は失敗。

初期化：$p \neq q, p \not> q, q \not> p$ なる $p \approx q \in E$ が存在したときは失敗。そうでなければ $i := 0$ 及び $R_0 := \{p \Rightarrow q \mid p \approx q \in E \cup E^{-1}, p > q\}$。

repeat　$R_{i+1} := R_i$

　forall $\langle p, q \rangle \in \mathrm{CP}(R_i)$ **do**

　　(a) $p\downarrow, q\downarrow$ を求める，それを各々 $\hat{p}, \hat{q}$ とする；

　　(b) $\hat{p} \neq \hat{q}, \hat{p} \not> \hat{q}, \hat{q} \not> \hat{p}$ であれば失敗して停止。

　　(c) $\hat{p} > \hat{q}$ ならば $R_{i+1} := R_{i+1} \cup \{\hat{p} \Rightarrow \hat{q}\}$;

　　(d) $\hat{q} > \hat{p}$ ならば $R_{i+1} := \mathrm{R}_{\mathrm{i+1}} \cup \{\hat{q} \Rightarrow \hat{p}\}$;

　od

　$i := i + 1$

until　$R_{i+1} == R_i$;

output　R_i;

↘ ［10］に從って「要對」とする。項書換系の完備化手續きにとって最も重要な概念である要對の定義を文獻［15］では「ある項が書き換え可能なとき，その書き換え對象となる部分の部分項も書き換え可能なとき，この 2 つの書き換えを『危險對（critical pairs)』と呼ぶ。」とするが，誤解を招く表現であるように思う。要對は二通りに書換可能な項を問題としている譯ではなく，（自分自身も含め）二つの書換規則の左邊に對する重像に對して定義されていること（特 q が變數でなく，最汎單化子が絡んでいることが）が重要である。

11） 完備化手續きは何か安定な整礎順序が與えられたことを前提として記述される。文獻［15］にはそのような整礎順序の記述が見られないが，他の考察によって停止性を假定しているのかもしれない。

3　包攝規準の形式的表現

3.1　包攝規準による等號系

文獻［15］は漢字學や特に包攝規準に興味をもつ讀者のために，數學的な説明を避けて，なるべく直感的な説明を與えようと試みている。從って，對象とする等號系，項書換系の定義を與えていない。これは却って本質的なところを見逃す結果になっているように思う。本節では前節の準備に基づいて，JIS の包攝規準を代數的に記述することを試みる[12]。

まず言語，卽ち定數を含む關數記號の集合を定義する。それは包攝規準に含まれる漢字部品と IDC（Ideographic Description Characters）から構成される。

・$\mathcal{F}_0$は包攝規準に含まれる漢字部品からなる。これらの定數記號を $a, a_1, a_2, \ldots, b, b_1, b_2 \ldots$などで示す。

・$\mathcal{F}_2$は二項 IDC からなる。卽ち $\mathcal{F}_2=\{$ ⿰, ⿱, ⿴, ⿵, ⿶, ⿷, ⿸, ⿹, ⿺, ⿻ $\}$とする。これらの二項 IDC を $C, C_1, C_2, \ldots$などで示す。

・$\mathcal{F}_3$は三項 IDC からなる。$\mathcal{F}_3=\{$ ⿲, ⿳ $\}$とする。これらの三項 IDC を $D, D_1, D_2, \ldots$などで示す。

加算無限個の變數の集合 X と關數記號の集合 $\mathcal{F}$ から構成される項集合 $T(\mathcal{F}, X)$ に含まれる項をここでは漢字構造記述と呼ぶことにする[13]。但し，IDS の記法に合わせて，例えば，二項 IDC C を使った漢字構造記述を $C(x, y)$ ではなく Cxy と書くことにする。また，關數記號のアリティが分かりやすいように適宜括弧を加える。

漢字構造記述についての等號系は，包攝規準を等式としたものから構成される。JIS の包攝規準の場合，これらの等式は以下の三つの型に分類される：

（Ⅰ）$a \approx b,$

（Ⅱ）$C_1ax \approx C_2ax,$

12） 本稿では JIS X 0208, JIS X 0213 で提出された包攝規準を「JIS の包攝規準」と呼び，本稿で使用する規準の番號もこちらの規格を參照する。また，その分析の多くは文獻［15］を參考にした。但し，本稿の目的は JIS の包攝規準に形式的記述を與えることにある譯ではなく，漢字構造の代數的記述の一例として，包攝規準を取り上げるのみである。

13） 文獻［15］では漢字構造記述と漢字（文字）の關係を何處で扱っているのかが分かり難い。文字體系に必須の漢字があるのであれば，F_0にそのための記號を導入し，その文字に關する關係を等式として E に追加する必要がある。そのような必要がなければモデルとして定義するのが自然である。

（Ⅲ）$C_1a(C_2xy)\approx C_2(C_1ax)y.$

JIS の包攝規準を具體的に示すと，Ⅱ型は 180 番⿰ ax≈⿺ ax（ここで a は「麥」）のみであり，III 型は 179 番⿱(⿸ bx)y≈⿸ b(⿱ xy)（ここで b は「厂」）のみである。他の全ての包攝規準はⅠ型に屬す。Ⅰ型に屬す包攝規準のうちで漢字部品が文脈に依存していて，兩邊に同一の IDC C を含む等式 $Cax\approx Cbx$ であると解釋すべき場合には，$f_a=\lambda x.\ Cax, f_b=\lambda x.\ Cbx$ に對應する一項の關數記號を新たに導入する必要がある。ここではこれらを除外して考えておく。

3.2　包攝規準による項書換系

これまでに定義した包攝規準による等號系から，完備化手續きを利用して完備な項書換系を構成しよう。

まず關數記號に順序を決める。Ⅰ型の包攝規準 $a\approx b$ から任意の方向に $a>b$ とする[14)]。Ⅱ型の包攝規準は⿰ $ax\approx$⿺ ax であり，⿰>⿺とする。Ⅲ型の包攝規準は⿸ b(⿱ xy)≈⿱(⿸ bx)y であり，⿸>⿱とする。この關數記號における順序 > から辭書式部分項順序を使って，各型の等式の兩邊には以下のように順序 ≫ が付く[15)]。

（Ⅰ）$a\gg b,$

（Ⅱ）$C_1>C_2$のとき $C_1ax\gg C_2ax,$

（Ⅲ）$C_1>C_2$のとき $C_1a(C_2xy)\gg C_2(C_1ax)y.$

これまでの準備に基づいて，包攝規準の等號系 E から完備な項書換系を構成しよう。まず定義 18 の完備化手續きの初期化として，等號系 E と上で決めた順序から R_0を與える。即ち $R_0:=\{p\Rightarrow q\mid p\approx q\in E\cup E^{-1}, p\gg q\}$ である。このとき書換規則の左邊は全て重像を作らないため $\mathrm{CP}(R_0)$ は空である。結局 R_0が求める完備な項書換系になる。包攝規準による等式は適切な順序を定義すると，要對を生成することなく，始めから完備な項書換系になっていた譯である[16)]。これは JIS の包攝規準の場合であるが，他の包攝規準を考えた場合，このようにうまくいくとは限らない。その包攝規準に適した手法が必要で

14）　Ⅰ型の包攝規準が三つ以上の漢字部品を包攝している場合，例えば a, b, c を包攝している場合も適當に $a>b>c$ とする。これはどの漢字部品が主になるかということを示すのではなく，項書換系のための機械的な重み付けである。

15）　Ⅲ型の場合は直ぐには導出できないが，辭書式部分項順序の定義を再歸的に使用して，$C_1a(C_2xy)>C_1ax, y$ であることから導かれる。

16）　全ての書換規則の左邊が二度以上の變數の出現をもたないとき，その項書換系は左線形 left-linear であるという。直交 orthogonal（左線形で要對のない）項書換系は合流性をもつことが知られている［2］。JIS の包攝規準による項書換系はこのような場合に相當する。

ある。

3.3 形式的體系の例として

これまでの議論を踏まえて，包攝規準の形式的表現について検討しよう。包攝規準は元々，符號化文字集合のための規則であり，文字コードの規格ごとに差異が存在する[17)]。從って，以下のように單純化して考えることにする

符號化文字集合は對象となる文字集合の各要素に一意の符號を割當てる。包攝とは，文字集合の複數の要素を一つの文字コード番號で符號化することを言い，包攝規準はそのための規則をさす。文字集合は文字をどのように抽象化するかによって決まるので，包攝の定義はその文字の抽象化を根據に定められる[18)]。卽ち，包攝は對象となる文字集合とその符號化に依存している。

これまでの本稿の定式化は文字の構造，包攝という概念を文字集合から切り離して，漢字構造を記述する形式的體系にのみ依存した概念とすることができた。漢字構造の代數的記述では，文字の構造や包攝は等式論理という形式言語によって議論し，様々な文字集合はモデルとして議論できるようになる。但し，本稿の例では文字コードの包攝規準も扱えるように，言語記號として漢字部品と Ideographic Description Characters（IDC）を用いるように制限を加えただけである[19)]。

Ideographic Description Sequences（IDS）は文字集合の要素である漢字の構造を表記する方法であるが，ここでの扱いは寧ろ形式言語での扱いに移行している。そして IDC は文字集合への關數として解釋され，JIS の包攝規準は JIS で想定された文字集合のみに限定される譯でなく，様々な文字集合，及びその上の解釋に適應される。

以上の漢字構造の代數的記述の特徴を踏まえて，この定式化の問題點と課題を幾つか述べよう。

包攝規準の適用除外　包攝規準の適用除外について，漢字「卷」「巻」の例で述べる。JIS 包攝規準の 14 番は等式龹≈𢍏，70 番は等式 己≈㔾で表現される。二つの等式から等式論理によって，等式 ⿱龹己≈⿱𢍏㔾が導かれる。しかし，包攝規準の適用除外を前提と

17）　例えば，Unicode の規格については文獻［1］などを參照されたい。

18）　文字コードの規格では包攝を定義するために，「字體」「字形」などの用語を使用するが，形式的體系として概念化するまでには至っていない。

19）　包攝規準を扱わないのであれば，漢字構造の記述に他の關數記號を使用することも可能である。これについては次節で述べる。

すると「巻」と「卷」は包攝できない。等式論理では部分項が包攝されるのであれば，項の全體も包攝されるべきであるが，⿱龹己≇⿱龹㔾であるということは，「巻」「卷」の構造記述は代入に關して閉じていない，或いは文脈と兩立しないため，等號系，項書換系を構成できないことになる[20]。

漢字の部品　言語記號を決めたとき，それによって構成される構造記述が漢字ではない場合がある。漢字構造記述には部品をIDCによって幾つでも合成できるが，それらの殆どが漢字に對應しないであろう。また，漢字の構造記述において，そのある部分が漢字ではない場合も存在する。上述の包攝規準の適用除外を含めて，漢字構造記述においては漢字であること，漢字の部品であることを形式化する必要がある。これには形だけでなく，音・義の形式化が課題となる。

完備化手續きの形式化　様々な包攝規準を等號系として形式化し，それを項書換系として利用するためには，完備化手續きの形式化を考慮すべきであったことを最後に付け加えておく。

4　應用について

これまで見てきたように，漢字構造を記述する形式的體系は言語記號を定義する部分（言語，シグニチャ），漢字構造に對する拘束條件を決める等號系（論理系などの統語論的な部分），それを充たす具體例である代數（意味を決めるモデル論的な部分）などから構成される[21]。

漢字構造記述の代數は，構造を記述する演算をもつ漢字集合である。様々な文字集合はその素材であって，構造を記述する演算を定義することで漢字集合としての代數を構成できる。前節の議論では關數記號としてIDCを想定していたが，必ずしも字體や字形を記述するものに限らず，解字などの情報も含むものとする。この枠組は漢字學の様々な場面で利用できる可能性がある。この漢字構造の形式的記述の利用方法について，以下で少し述べてみたい。

20）　包攝規準の適用除外について，文獻［15］では完備な項書換系でないことと關連付けているが，これは誤解を招きかねない。

21）　文獻［11］ではこれらの構成要素を一つに纏めてinstitutionとして提出する。Institution理論について，詳しくは文獻［4］を參照されたい。

1. 符號化文字集合と包攝規準
   様々な包攝規準を適用して，符號化文字集合を新たな漢字構造記述の代數として扱うことができる。また，モデルとしての文字集合に對して，包攝規準の檢證に利用できる。その典型例を文獻［8］で見ることができる。
2. 「文字編」の枠組の形式的な表現
   竹簡などの古文字資料に對して，その文字を整理するための枠組を提供する。一般的な「文字編」は說文の部首體系により整理されているが，それ以外の體系に對しても同様に形式的體系を與えることが可能である。

以上は言語，等號系，代數を一つの文字體系を記述するために使用した例であるが，文獻［11］では複數の文字體系を利用する爲の枠組も提供する。即ち，複數の文字體系の統合を記述することである。例えば，一つの漢字に注目しても，どの文字體系の中で考えるかということに依存して解釋が異なる譯で，從って，漢字單獨で記述するのではなく，體系の中での複數の文字の張り合いを記述する必要がある。詳しくは文獻［11］を參照してほしい。この枠組を使用して，以下の課題への利用を考えている。いずれも複數の文字體系の統合を形式化できることが必須である。

1. 統合的な字書
   文獻［13］では編集方針が異なる字書を混合した場合に生じる問題について議論している。ここで述べた複數の文字體系の統合を形式化することの檢討を提案したい。
2. 文字學的な文獻の翻刻
   文獻［12］では『說文解字注』の翻刻について考察した。そこには說文の文字概念，引用文獻の文字概念，段氏が使用する文字概念など，様々な漢字を翻刻する必要がある。

5 お わ り に

最後 Granger の形式的思考に戻って，漢字構造の代數的記述について議論しよう。目に見える視覺的現象である漢字の形は，屢々既知の對象であると考えられがちであり，知覺的對象と形式的對象を同じものとして扱う傾向が非常に強い。しかし，本稿の簡單な形式的扱いでも明らかになったように，視覺的對象である漢字の形でさえ，形式的體系によってその內容が與えられることが理解されるであろう。更に漢字であること，漢字部品であることを內容とする形式化が次ぎに目指すべき課題である。節 3 で述べた漢

字構造の形式的記述の課題は，漢字の形音義を含めた統合的な形式化によって取組む問題であろう。

漢字情報の形式化について文献［11］で議論したが，本稿で行なった漢字構造を代數的に記述する方法はその一部と捉えることができる。そこで用いた Institution 理論は様々な意味論的な體系を形式化したものであり，形式的思考における「主題化」と解釋できる。文字體系の記述，及びその統合を形式化するということは，人文學の對象に對する形式的思考を押し進めることであり，また，多くの可能性を祕めていると考えている。

謝　辭

Christian Wittern 先生には本稿の草稿に對して，多くの有益な助言を戴きました。ここに感謝の言葉を述べさせて頂きます。多くの方々からは感謝し盡くせない援助を戴きました。また，いつも支えてくれる妻留美と娘に感謝します。

参 考 文 獻

［1］J. D. Allen et al. eds.: The Unicode Standard Version 6.0 — Core Specification, the Unicode Consortium, 2011.

［2］F. Baader, T. Nipkow: *Term Rewriting and All That*, Cambridge University Press, 1998.

［3］S. Burris, H. P. Sankappanavar: *A Course in Universal Algebra*, Springer, 1981.

［4］R. Diaconescu: *Institution-independent Model Theory*, Birkhäuser, 2008.

［5］G.-G. Granger: The Notion of Formal Content, *Social Research* 49(2): 359-381, 1982.

［6］G.-G. Granger: *Pour la Connaissance Philoso-phique*, Éditions Odile Jacob, 1988.（邦譯 植木哲也譯：『哲學的認識のために』，法政大學出版局，1996.）

［7］G. Grätzer: *Universal Algebra*, 2nd ed., Springer, 1979.

［8］川幡太一：IDS による UCS 漢字の「同一性」の判定手法, 東洋學へのコンピューター利用 17: 105-119, 2006.

［9］近藤和敬：グランジェの科学認識論 ——「操作－對象の雙對性」,「形式的內容」,「記號的宇宙」，所收 金森修編：『エピステモロジー —— 20 世紀のフランス科學思想史』，慶應義塾大學出版會，2013.

［10］坂井公：Knuth-Bendix の完備化手續きとその應用，コンピュータソフトウェア 4(1): 2-22, 1987.

［11］白須裕之：文字體系の統合による漢字情報の形式化 ——『說文解字注』における音注を事例として，情報處理學會「人文科學とコンピュータシンポジウム」論文集，209-216, 2013.

［12］白須裕之：漢籍の電子的な翻刻について ——『說文解字注』の Unicode 轉寫を事例として，東方學報 89: 211-238, 京都大學人文科學研究所 2014.

［13］鈴木俊哉：文字分類方式の變更が字形に及ぼす影響，情報處理學會研究報告 2012-DD-86(3),

2012.
[14] 外山芳人：完備化による等式證明，人工知能學會誌 16(5) : 668-674, 2001.
[15] 守岡知彦：項書き換え系を用いた漢字字體の包攝規準の形式化の試み，情報處理學會論文誌 59 (2) : 332-340, 2018.