

4J-4

日本語文書作成支援環境IDEにおける
文書の構造的作成のためのエディタ山口 浩司 宇津宮 孝一
(九州大学 総合理工学研究科)

1 はじめに

理想的な文書処理システムは、一貫した概念に基づいて構築され、一様な利用者インタフェースを提供しなければならない。さらに、文書のライフサイクルのすべての段階で利用者を支援する必要がある。文書には、計算機のソフトウェアの場合と同様に「作成」、「流通」、「消滅」という3つの段階に大別されるライフサイクルが存在する。我々は文書処理支援環境構築の端緒として、文書のライフサイクルにおける文書の作成過程に注目して、日本語文書の構造的作成を支援する環境の構築を試みた。

2 文書の構造的作成

文書の作成は、通常、文、段落、図、表のような個々の具体的内容の記述とそれらの間の論理的な流れの決定という2つの側面からなされる。その作成過程は書き手のスタイルによって異なるが、代表的な作成法として2つの作成法が考えられる。1つは、最初に文書全体の論理的な流れを決定し、それに基づいて文章等を書き進めるものである(下降型)。これは、特に論理性の高い科学技術論文などの作成に望まれる作成法である。もちろん、その論理的な構成が作成途中に変更されることも少なくない。もう1つは、最初に文書の素材となる文章や図表を揃えておき、それらを論理的に組み合わせながら文書を作成するものである(上昇型)。この作成法は、書き手が過去に書いた文章や図表あるいは他の文献や資料からの引用を行いながら新たな文書を作成する場合に用いられる。文書の共同作成を行う場合もこのような作成法を用いることが多い。

これまでに提供された文書処理システムの多くは、上述した文書編集の2つの側面のうち、個々の具体的内容の編集に重点が置かれている。文章に関しては文字列としての処理が中心である。図表の作成に関しても、さまざまな道具が提供され、高品質な図表が作成できるようになってきた。しかし、文章と図表との統合は、一般にレイアウトという物理的な側面での編集によって行われ、その間の参照関係のような論理的な情報についてはほとんど活用されていない。文書の構造という論理的な側面に注目して文書の作成やその後の処理を行えば、より効果的に文書の作成を行うこと

ができる。それはまた流通段階においても有効である。

構造が構文規則に従うようなプログラムに対して、文書の論理的な構造は、文書の用途(論文であるとか、手紙であるとか)や書き手のスタイルに依存している。それ以下の部分は自然言語としての文章であり、その構造を文書の論理的な構造と同一の概念で取り扱うことはできない。また、文書の構造を扱う場合、文書の論理的な単位としては、章、節、段落などが適当であり、それらの構成関係も、できるだけ単純な構造として表せば、書き手による構成の編集も容易になる。文書における論理的な単位は、最終的に文書が消滅される際に、その外観や体裁として特定の見え方の属性をもつことが多い。文書には論理的なまとまりごとに宣言的に記述できるような書式もある。文書の構造情報は流通段階にも有用である。

このような考察を基に、我々は文書の論理的な構造の把握や変更を容易にし、また書式化や管理のための情報を扱う文書エディタを実現した。

3 データ構造と基本機能

データ構造は、基本的には階層ファイルシステムとして実現する。文書はその論理的な構成に基づいて、図1に示すような木構造(文書構成木)で表される。文書構成木の部分木が文書の論理的な構成要素を表し、終端ノードが具体対象(テキストなど)に相当する。一方、文中で参照される具体対象を表すノードは、その種類ごとに*図、*表、*文献などの各ノードの下に

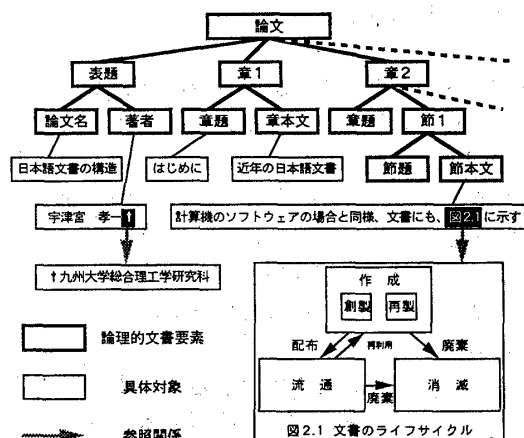


図1 文書の論理的な構造と参照関係

まとめる。文中からの参照関係は、その間を双方向にポインタで結合することによって表現される。これによって、1つの具体対象（図、表など）に対する文中の複数箇所からの参照も可能になるとともに、同種のノードが文書全体に散在するのを防ぐ。また、個々のノードの編集を文中での参照状態とは独立に行うことも可能である。文書構成木上の各ノードには、具体対象に対する書式や部分木全体に対する書式も格納する。

システムは、このような木構造を取り扱うエディタと、文書構成木の各具体対象の入力・編集を行うための各種エディタ（日本語テキストエディタ、図表エディタなど）を用いて実現される。木構造エディタは、画面上に表現された木構造に対して、ノードの生成、削除、分割、併合と部分木の移動、複写を行う機能を基本的にもつ。これらの機能を用いて、文書構成木の編集を行う。

4 文書の編集

本システムでは、文書の構成（文書構成規則）や構成要素ごとの具体対象の種類（具体対象生成記述）とその書式（書式記述）が定義された文書クラスを用意している^[1]。もちろん、利用者が定義することもできる。各文書はある文書クラスの1つのインスタンスとして生成される。

文書クラスのインスタンスとして生成される文書には、文書構成木の初期値として、その文書クラスの文書構成規則が与えられる。その文書構成木にノードに対する操作を対話的に行うことによって、最終的な構成へと具体化が行われる。具体化された文書構成木の各ノード（文書構成要素）に対しては、それぞれの具体対象生成記述に基づいた具体対象の入力や書式の定義が行われる。構成の編集は、具体対象の入力や書式の定義を行いながら、構成要素の追加や分割、併合という形で任意に行うことができる。

5 プロトタイプIDE-1

システムの有効性の検証と効果的な利用者インタフェースの確立を早期に行うために、プロトタイプIDE-1を試作した。支援対象としては、計算機を道具として利用する研究者の定型的文書（論文、手紙、通知等）が中心である。これまでに、汎用計算機のTSS上でフルスクリーン型日本語端末を使用して実現しており、日本語テキストを中心に扱っている。現在、小型計算機上でも実現中である。

構造的に作成された文書は、さまざまな形での閲覧が可能である。たとえば、文書構成木を基に、特定の章、節のみの閲覧（書式化）や再利用が行える。また、文書中で使用されている図表のみの閲覧などにも対応できる。このように、構造的に作成がなされた文書の構造情報を作成以降の段階においても活用することによって、よりよい文書処理支援環境が実現できる。

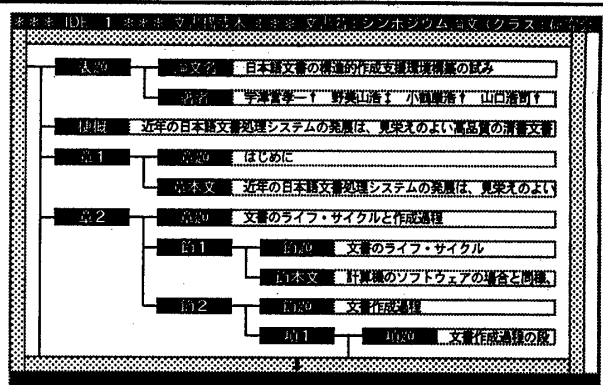


図2 文書構成木表示画面

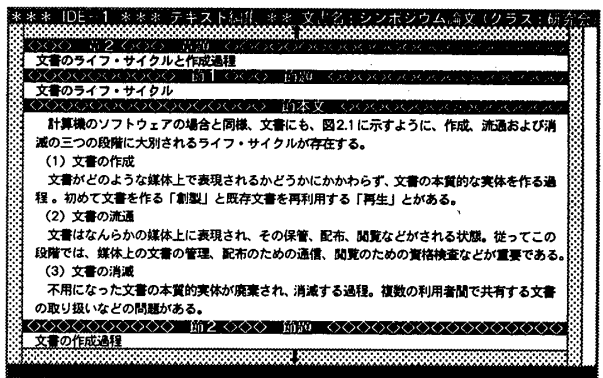


図3 テキスト編集画面

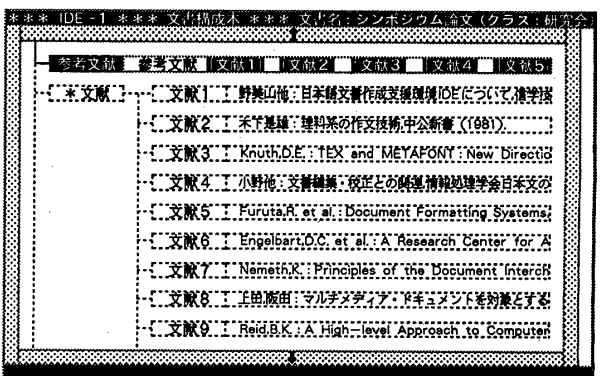


図4 文献参照の例（文書構成木表示）

6 おわりに

我々は、文書の構造という形式的な側面に着目して文書の構造的作成を支援するエディタの設計を行い、そのプロトタイプを作成した。ここでは、文書の構造を木構造として表現し、文書の論理的な構成の把握と編集を容易にした。今後は、文書エディタとしての利用者インタフェースの改善を図るとともに、文書の管理や通信という流通段階をも含めた統合文書処理支援環境の構築を進めてゆく計画である。

参考文献

[1] 宇津宮他：日本語文書の構造的作成支援環境構築の試み，情報処理学会「日本語文書の入力と編集」シンポジウム論文集，pp.93-104（1985.11）．他