

# プログラミング言語教科書の動的構築

花川賢治

大阪府立工業高等専門学校 電子情報工学科

## 1 はじめに

従来から、普遍性の高い情報を自然言語で表現して配布する方法として「書物」という形態が使われてきた。しかしながら、書物による情報の流通システムにおいては、異なった著者が同一の情報を記述するという冗長性の問題と、誤った情報を記述するという品質の問題が必然的に存在する。筆者らはこの問題を解決するために、図1に示すような動的文書システムを提案してきた。このシステムは、ネットワーク上で共有される文書レポジトリに自然言語で記述した情報の断片である文書部品を多数格納し、ユーザから要求があったときに、必要な文書部品を取出し一つの文書を構成し提示する。文書レポジトリに格納される文書部品は多くの専門家が真であることを合意したものに限り、冗長性がないように管理される。

動的文書システムにおいては、文書部品は文脈非依存性が要求される。ここでの、文脈非依存性とは、外部環境とは無関係に必ず記述の解釈が一意であることを指す。情報を文脈非依存な文書部品で表現できるかどうかは、対象となる情報の内容に依存する。情報が普遍的、客観的な事実であれば、文書部品を文脈非依存にすることができると考えられる。[1]では省略と指示照応に関係した部分を補足するという軽微な修正で、人間が書いた文書を文単位に分解したものを文脈非依存な記述に書き換えることができることを示した。

人間にとって読みやすくわかりやすい文書を自動構築するためには、実際に人間が書いた文書に内在する構造をモデル化したものに基づく必要がある。[1]では実際の文書を調査して、そこにみられるトップダウン構造と概念強化構造に従い、いったん分解した文書を再構成し、それが人間にとってリーズナブルであることを確認した。

以上の研究を踏まえ、本研究では、文書一般の文脈非依存化、構造のモデル化の可能性をより詳しく調べる。文書部品の文脈非依存化と構造化の問題は、文書

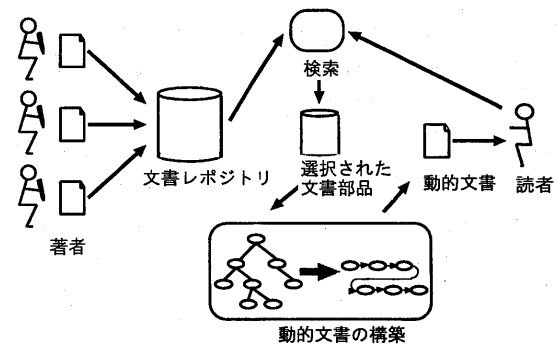


図1: 動的文書システム

の内容、利用目的、利用形態に強く依存すると考えられるので、学校での教育現場で有用な対話的な教育支援システムを構築するという具体的な課題を設定した上で、この問題を考察する。さらに実用的なシステムを運用する過程で人間の主観的評価を得ることも意図している。

## 2 動的文書の教科書への応用

一般に学校教育の特徴として、扱う内容が比較的普遍性が高い、すなわち文脈依存性が低い傾向がある。しかしながら、実際の教科書には、文脈非依存化が困難な部分が多くみられる。たとえば、C言語の教科書では、リファレンスマニュアルの部分（[2]の付録A）は文脈非依存化と構造のモデル化が容易であるが、チュートリアル部分（[2]の本文）は非常に困難である。チュートリアル部分の部分はプログラム例への参照と、推論の展開が多く含まれ、分解すると各部分は意味をなさなくなる。

この現象を教科書で提示される情報を以下のように分類して考察する。

1. 一般的な事実の記述
2. 具体例
3. 上の二つを対応づける説明

一般的な事実の記述は、学習者が理解すべき内容そのものであり、教科書の内容として必須の要素である。しかし、一般的な事実だけを示されても、学習者はそれを理解するのは困難であり、補助的な手段として、具体例を示す必要がある。さらに、学習者が自力で両者を対応づけることができないならば、両者を対応づける説明を示す必要がある。

一般的な事実の記述は文脈非依存化と構造のモデル化が容易である。それに対し、具体例は、教師が自由に導入できるもので本来普遍性が低い。また、一般に構成要素の独立性が低い。両者を対応づける説明は、人間の知的な活動により両者から生成可能なものであり、その記述の構造は推論の構造に一致し、要素間の関係は複雑である。リファレンスマニュアルは一般的な事実の記述のみから構成されるのに対し、チュートリアルは具体例と両者を対応づける説明が含まれる。そのためチュートリアルの部分は文脈非依存化が困難になると理解できる。

動的文書の教科書への応用として、一般的事実を文書部品として文書レポジトリに格納しておき、与えられた具体例を文書部品を使って説明するシステムが考えられる。このようなシステムでは、文書部品は文脈非依存であるのに対し、生成される文書には文脈依存性が付加されることになる。

### 3 動的 C 言語教科書

ここでは、現在開発中の動的 C 言語教科書、UNI-C について述べる。

C 言語教科書のチュートリアルでは、プログラム例の一部分について、それが何であるか、そのことが言語仕様からどのように導かれるのか記述した部分が大部分を占める。UNI-C の基本的な機能は、プログラム例が与えられたときに、自動的に C 言語仕様記述の断片を加工して説明文書を生成することである。

C 言語仕様記述に現れる一般名詞 (構文カテゴリ、型、スコープ、オブジェクト等の種類など) が述語名、プログラム例を構成する構文解析上の記号に対応した部分文字列が個体であると考え、プログラムの一部分を引数として与えると述語の真理値が決定できる。また、C 言語仕様記述の断片は一階述語論理式で表すことができ、述語が真であるときには、それを証明することができる。この考え方は、広く知られているプログラミング言語の字句解析、構文解析を推論処理とする考え方を拡大したものとみることができる。

UNI-C は C 言語仕様の自然言語での記述とそれに対応した一階述語論理の記述の両方を格納し、前者を教科書を構成する部品、後者をプログラムを解析する推論処理に用いる。UNI-C は以下のプロセスを対話的に実行する。

1. 教師または学習者からプログラム例を得る。
2. プログラム例を解析し、それに含まれる部分文字列を引数とする真である fact の集合を生成する。そのとき C 言語仕様に対応した一階述語論理式を使った推論を行う。
3. 学習者に興味対象であるプログラムの部分や概念を表すキーワードを入力させる。
4. 学習者の興味に関連のありそうな fact とそれを導いた推論の過程を抽出する。
5. 推論の過程に対応する文書部品から説明文書を構成し学習者に提示する。

たとえば、UNI-C に "hello" を出力する C プログラムを入力し、部分文字列 "printf("hello\n")" を指示して説明を求めると、それが式の正しい形式であり、標準関数である printf の引数の条件に合致すること、printf が識別子、"hello\n" が文字列リテラルの条件を満たすことなど、プログラム例の妥当性を説明する文書を得ることができる。その説明を読むことにより学習者は C 言語仕様の理解を深めることになる。

### 4 おわりに

動的文書システムは情報流通の効率と品質を飛躍的に向上させる可能性を持っている。学校教育は情報流通の場の一つであり、本研究はそこで動的文書を応用することを目的としている。教科書は表面上は文脈依存性が高いように見える。しかし、核になる部分は普遍的な情報であり、動的文書システムで扱うことが可能である。また、それを使って個別に与えられる具体例の説明を生成することも可能である。

### 参考文献

- [1] K.Hanakawa, H.Takada and T.Nishida: Construction of a Dynamic Document Using Context-Free Pieces, Proc. the 11th international conference on software engineering & knowledge engineering, pp 212-216, (1999)
- [2] B.W.Kernighan, D.M.Ritchie: The C Programming Language, Second Edition, (1988)