

TM-0166

プログラム自動生成

岩元莞二、西谷泰昭、和田 孝
(日本電気)

April, 1986

©1986, ICOT

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

プログラム自動生成

Automatic Program Generation

岩元 寛二*

西谷 泰昭*

和田 孝*

Kanji Iwamoto

Yasuaki Nishitani

Takashi Wada

* ソフトウェア生産技術研究所 生産技術開発部

Software Product Engineering Laboratory

Abstract

Usually, module specifications given to programmers are written in the combination of natural language (Japanese), figures, and tables. This paper introduces experimental system PGEN-1 which understands such specifications and automatically generates computer programs. Although restricted number of vocabulary is allowed in the current system because of ^{the} experimental small scale dictionary, the system can understand, to an extent, semantically ambiguous and abbreviated expressions and produce programs as the result.

要旨

プログラマ向けに書かれたモジュール仕様書は、日本語の文章、図、表の組合せで記述されている場合が多い。本稿では、このような自然な形式で書かれたモジュール仕様書を理解して自動的にプログラムを生成する実験システムを用いるので、使用できる語いは制限されまいが、意味的なあいまいさや省略のある表現もある程度理解してプログラムを生成することはできます。

1. よゑかき

プログラムを作成するときは、普通は日本語の文章や図表で処理内容を記述してから、それをプログラムの言語で書き直すという手順を踏んでいきます。特に、プログラムの設計とプログラムの製造（プログラミング）の間を明確に別工程として区切り、担当者も別にしていう場合は、設計者とプログラマのコミュニケーションのため詳細なモジュール（プログラム）仕様書が利用されます。

プログラム自動生成実験システムPGE-1は、プログラマ向けのモジュール仕様書を理解し、プログラム言語COBOLによるプログラムを自動的に生成する実験システムです。

PGE-1の間接では仕様書の理解が主な技術課題となっており自動翻訳と類似していましたが、適用対象をモジュール仕様書に限定して意味解析を十分に行う点で違いました。例えば、複数の解釈の可能性があり異なる表現や

省略した表現があっても、業務用語やプログラミングに関する知識と人間の経験的解釈規則を利用して正確な解釈を行うように努めます。

この度開発したPGE N-1の概要について、次に紹介します。

2. システム構成と機能概要

PGE N-1は図1のようなシステム構成をとしています。システムの流れは、~~ざきさだ~~自由度の高い日本語の文章と図表で書かれたモジュール仕様書を理解してプログラムを自動生成することです。ただ実験システムですが、基本的に以下の特徴を持っています。

- ① 日本語の文法上の制限を意識させない。
- ② 業務用語を用いた仕様記述と可能にする。
- ③ 使用できる用言（動詞、助動詞）は制限される。
- ④ ある程度のあいまいさや省略のある文を理解できる。
- ⑤ 日本語の文章と図表を組合せて意味を理解する。

以上の特徴により、使用でさき単語に制限はありませうが、
モジュール仕様書とせうには十分な自由度がありませう。 ← 図1

ただし、本システムを利用してプログラムを作成する前に
に準備作業が必要です。モジュール仕様書の中でデータタ
名に付添する業務用語が使われませうが、業務用語とデータ
の対応関係、及びデータ構造の設計結果は、辞書作成ツ
ルを使つて予め変数辞書とデータ型辞書の中に登録してお
くことが必要です。システムは対象ソフトウェアに固有な
これらの二つの辞書をもとに修正された構文解析用辞書と意
味処理用辞書を参照しながら、モジュール仕様書を理解し
てプログラムを生成します。そのときの処理の流れは図1
に示す通りです。

入力処理部は、仕様書作成用のエディタ機能と入力の構
造解析機能を兼ねています。日本文、表、図の構造解析結
果はそれを小エディタ、エディット、エディットといふネッ
トワークで表現させます。意味処理部はこの三種のネッ

トワークを統合して仕様書の意味を表すMネットを作成します。その後、Mネット上で冗長な情報^①の削除、意味的あいまいさの解消、省略語の補完などの意味処理を行^②、乙、プログラムの生成に通した形になるようにMネットを変形します。プログラムの生成はMネットをその構造に従^③、乙解釈しながらCOBOLのプログラムを生成します。

3. モジュール仕様の記述例

本システムが処理できるモジュールは様の記述例と図2と図3で示します。図2はプログラムの手続き部の処理内容も文章で記述したものであり、その中で参照される^④表と図が図3です。

尚、データ宣言部は変数辞書から自動生成されます。変数辞書に変数定義を登録するとその記述形式は文章形式ではありません。COBOLのデータ宣言形式に近いもので

トリ 文章の入力は漢字エディタを使用、て分ち書きで入力します。

次に記述例に従って本システムが処理できる仕様記述の特徴を説明します。

1) 節・段落構造

仕様記述は段落構造を持つといえます。節は節番号(「1」で囲まぬ)と節タイトル(省略可能)からなり、節ヘッダで始まり、節の本体に付いた節が合すものとします。(1)、(2)などであり、簡単な電文と記述可能です。

2) 節タイトル利用

節タイトルは通常はコメント扱いですが、節本体の記述中の情報不足を補う役割を果します。現在オミタムでは、データの修飾子を見つけてのみに利用していません。例えば、図2の「2」節と「3」節で「コ-ド数」という名詞が出てきますが、データ型積書の内容と節タイトルから、「2」節では、監査キ-コ-ドエリアの、「3」節では注意コ

ードエリアのサブファクトールDの名前で断ります。

3) 「以下」の使用

「以下を行なう」、「以下の処理を行なう」という表現で「以下」がどの範囲かを理解する必要がありません。オシロスコープはこれらのため、段落の構造下利用してります。

4) 処理条件の簡略表現

「2」節の「有帳のとき」は、正確には「処理区分が有帳のとき」ですが、データー型辞書を参照して「有帳」がコード表「処理区分」のコード値であることを知り、正しく解釈することが出来ます。

5) 主語や目的語の省略

「有帳のとき」でも主語が省略されていますが、[4]節の「22」のときは「主語である科目コードが省略され」、「N1」をセットでは目的語である摘要コード(に)が省略されています。この場合は文のパターンから省略語を補完出来ます。このほかにも、局所的な文脈が

2
省略語を補完します。例えば、「AとBを加え、Zに代入する」と「AとBの和をZに代入する」と変換してZに代入させるのは、さりさることであります。

6) 表の参照

「5」節では条件の一部を文章で述べ、残りの組合せ的な条件は表を参照してします。また、処理内容の大部分は文章で述べられていますが、どんな値をセットするかが表の中で示してします。本シラスは文章と表の意味を結合して完全な仕様として理解します。

7) 図の参照

「7」節と「8」節は図を参照した例です。同一図でもそれを参照していき文章が違っても図の意味は違ってもです。「7」節では条件を表すものとして理解し、「8」節では代入の関係を表すものとして理解します。また、図は完全なものではなくても、データの型情報を見ても足情報も補うことができます。

以上のほか、記述例にないような正しいな表現形式に表現もある程度処理できますが、詳細は省略します。

4. 仕様理解のための知識

PGE N-1 システムは仕様理解とプログラム生成のためにはさまざまな知識を使っています。それは、プログラムに組込みの知識と分離した知識に分けられます。プログラムに組込みの知識は主に日本語の文法に関する知識と、意味解析法に関する経典的知識です。プログラムから分離した知識は、PGE N-1の実行時には構文解析用辞書と変換処理用辞書に集約されますが、これは事前に変換ツールを使い、日本語マスタ辞書、データー型マスタ辞書、変数マスタ辞書、ハズデル(辞書)から生成可能です。

以上の知識の中でプログラム生成に特有のものは、デ-

タ型マスタ辞書、変数マスタ辞書、及びモデルズ。特にモデルは、対象ソフトウェアに依存しないもので、プログラムの構成要素（演算子，データ，式，関数，ルーチン構造，分岐構造，etc.）に関する知識を持つています。本システムはモデルに反しないように仕様を解釈します。

5. おまが

PTE-M-1は実験システムでより精密なモデルを実験用的小型模のモデルが、幾つかの例題について入力からプログラムの生成までを行えることを確認しました。今後はさらに能力の安定化とレベルアップのため研究が必要と考えられます。尚、本研究は、(財)新世代コンピュータ技術開発機構からの委託（発注3401-04号）の一環として行なわれています。

参考文献

- 1) 「PGEN-1: 仕様書理解をベースとしたプログラム自動生成システム」, 岩元覚二, ~~田中~~ 情報処理学会第32回全国大会.
- 2) 「数式も含む日本文の構文解析」, 和田孝. ~~母が~~ 同上

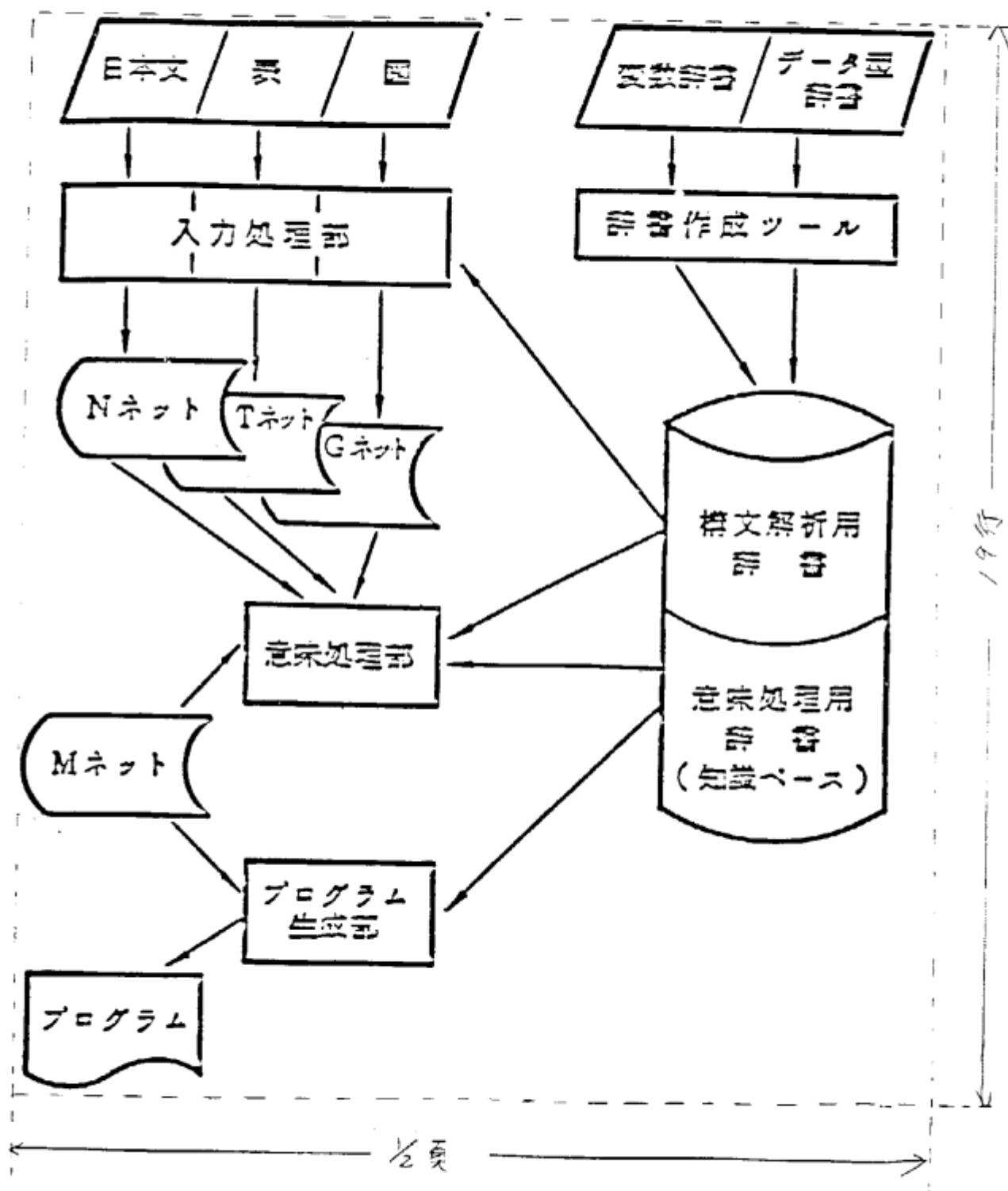


図1 PGEN-1 のシステム構成

- 198
- [1.] 入金総額の計算
金種が "00" でない入金明細について、以下を繰り返す。
 (1) 総額に入金額を加える。
 (2) 入出金口数をカウントアップする。
 - [2.] 監査キーコードエリアのセット
有帳のとき、以下の処理を行なう。
 [2. 1.]
 使用区分が1のとき、以下を行なう。
 (1) コード数をカウントアップする。
 (2) コード数が5以下なら、コード数番目のコードを "321" とする。
 0のとき、後送以外なら、エラーコードを "321" として、リターン。
 - [3.] 注意コードエリアのセット
有帳のとき、以下の処理を行なう。
 [3. 1.]
 使用区分が1のとき、以下を行なう。
 (1) コード数をカウントアップする。
 (2) コード数が5以下なら、コード数番目のコードを "321" とする。
 0のとき、後送以外なら、エラーコードを "321" としてリターン。
 - [4.]
科目コードが "20" のとき、摘要コードに "N0" をセットし、
"22" のときは、"N1" をセット。
 - [5.]
追残がマル限限度額より大きいとき、表2に従い摘要コードに値をセットしてリターンする。
 - [6.]
流動性預金のとき、表1の組み合わせ以外なら、エラーを409としてリターンする。
 - [7.]
図1のとき、リターンする。
 - [8.]
図1に従い、値をセット。

1/2頁

図2 日本語による仕様記述

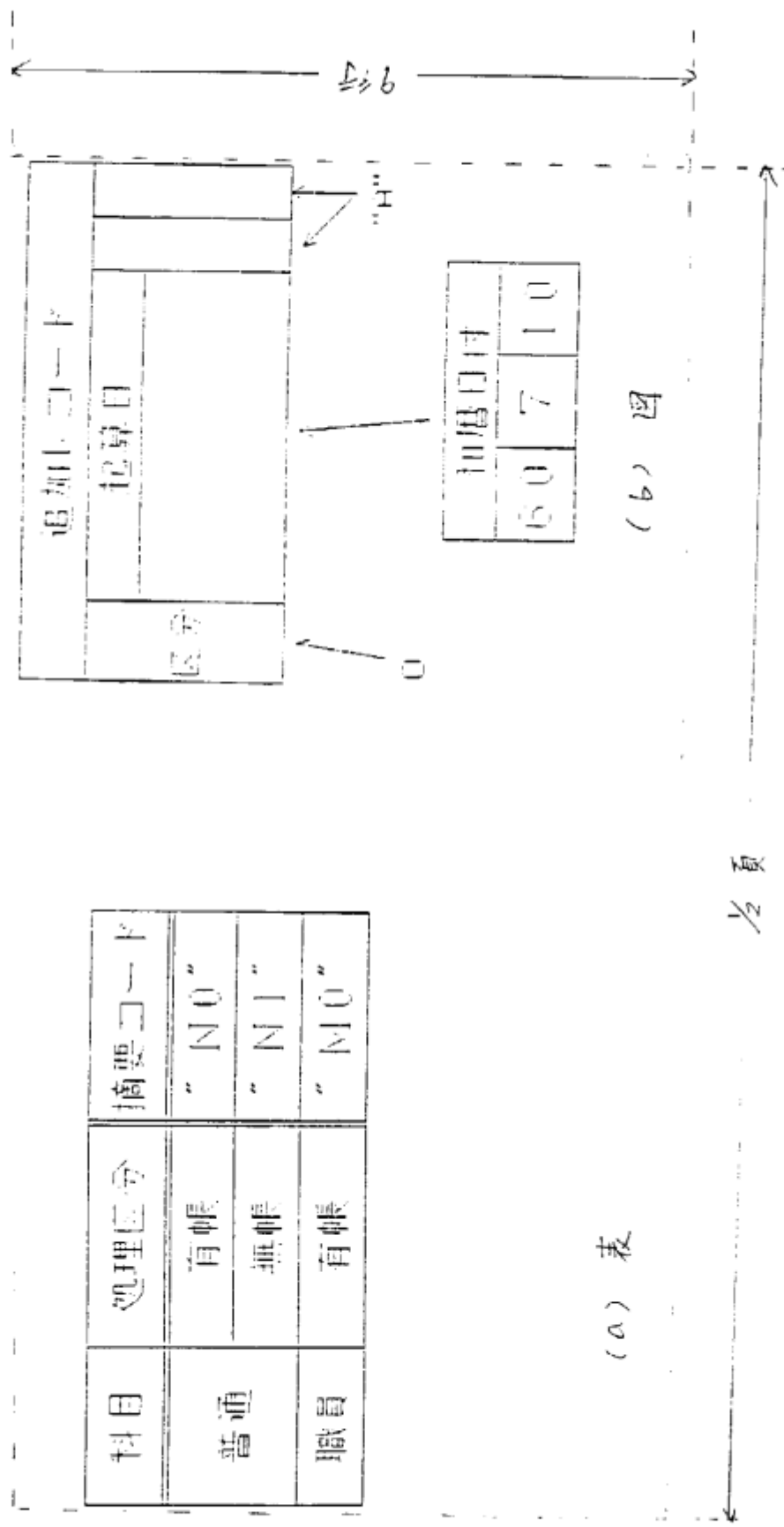


図3 表と図の例