

日本科学技術情報センターにおける文献データベースシステム

Nation-wide Online Information Retrieval System of Japan Information Center of Science and Technology

特殊法人日本科学技術情報センターでは、昭和51年から科学技術情報を中心とするオンライン文献検索サービスを行ってきたが、対象データベースの拡大、サービス機能の拡大を図るため、新たにJOIS-IIを開発し、昭和56年4月からサービスを開始した。

JOIS-IIは不特定多数の一般利用者を対象とした商用システムで、日本全国1,000のユーザーから約1,200万件のデータベースを検索することのできる大規模なオンライン情報検索システムである。本稿では、JOIS-IIのシステムの概要について述べる。

長谷川 昇* Noboru Hasegawa
諏訪 秀策* Shūsaku Suwa
田中 重康** Shigeyasu Tanaka
宮本 一治** Kazuharu Miyamoto

1 緒言

「情報爆発の時代」と言われている最近、科学技術の分野でも大量の情報が発生している。その数は科学技術関係刊行物の論文・記事、特許資料、技術レポート、その他の科学技術情報を合わせて世界で年間約400万件も発生していると言われており、その伸び率は年率約5%の勢いにあるという。こうなると研究者にとって大量の情報の中から必要とする情報を自力で探し出すことは至難の技となり、ましてや時々刻々発生する世界の新しい情報を漏れなく網羅的に探しだすことは不可能である。これにこたえるシステムとして、世界ではいろいろなオンライン文献検索システムが開発され、そのサービスに当たっているが、とりわけ、その代表的な商用サービスシステムとしては、米国ロッキード社のDIALOG、SDC社のORBIT、BRS社のBRSといったシステムがある。

一方、JICST(Japan Information Center of Science and Technology：特殊法人日本科学技術情報センター)は我が国の情報サービスの中核機関としての役割を果たすために、米国にみられるレベルのシステムを開発し、情報検索サービスを提供してきた。

特に、JOIS-II(JICST Online Information System II)がDIALOGやORBITと異なる点は、国内及び国外で発生する文献情報(JICST科学技術文献ファイル、JICST科学技術研究情報ファイル及びJICST国内医学文献ファイル)に日本語の抄録を付けてサービスを行なうことにある。実験用オンラインサービスシステムとして、昭和51年7月からサービスを開始した英字・片仮名情報ベースのJOIS-Iに続き、同54年2月から特定回線による漢字情報ベースのJOIS-Kの実験サービスを開始し、続いて同54年7月からはJOIS-I、JOIS-Kを統合し、更に機能拡大を図ったJOIS-IIの開発に着手し同56年3月に完成の上、同年4月からサービス開始となった。以下本稿では、JOIS-IIの開発方針、システムの概要、稼動評価及び今後の課題について述べる。

2 JOIS-IIの開発

2.1 開発の目的

JOIS-IIは、本格的な営業システムとして日本全国の一般

の研究者や技術者に、我が国はもとより世界各国の科学技術文献情報、研究案内情報の検索サービスを提供することを目的として開発された。

2.2 システム開発の基本方針

これまでのJOIS-I、JOIS-Kサービスでの実験結果の評価をもとに、下記のような基本方針を設定した。

- (1) データベースの量的・質的充実拡大
- (2) 会話機能の強化
 - (a) 検索機能の追加
 - (b) 質問登録機能の追加
 - (c) シソーラス(辞書)ファイルのサポート
 - (d) 複写受付機能の追加
 - (e) JICST資料所蔵目録ファイル参照機能のサポート
- (3) JOIS-I、JOIS-K及びバッチ検索システムの統合
- (4) 公衆漢字端末のサポート
- (5) 営業用システムとしての運用上の考慮(課金方式)
- (6) 接続可能端末数の増加(処理能力アップ、回線ネットワーク拡大)

3 システム概要

3.1 データベース

サービス対象データベースは、表1に示すとおりである。JOIS-IIは基本的には文献検索システムであるが、研究案内情報や文献検索の補助手段として用いる所蔵資料、用語、化合物などのファイルも文献類似の取扱いでサービスを行っている。

3.2 会話機能

オンライン端末から入力される利用者の検索質問に答えて、文献データベースの中から該当の文献情報を会話的に検索することにより、目的とする文献に絞り込んでゆくことを可能としている。その会話例は図1に示すとおりである。

会話コマンドは表2に示すとおりであるが、JOIS-IIは検索機能、検索補助機能、質問登録機能、会話補助機能などを追加し、コマンド数を倍増させた。特に質問登録機能は、利用者が必要とする文献テーマについて、キーワード、論理式

* 特殊法人日本科学技術情報センター ** ファコム・ハイタック株式会社

表1 JOIS-IIのサービス対象データベース 文献ファイルだけでなく、研究情報ファイル、サポートファイルとサービス範囲は広く、蓄積データ量も約1,200万件に及ぶ。

(a) データベース

	データベース名	収録期間	収録件数	対象分野	備 考
文 献 フ ァ イ ル	JICST 科学技術文献ファイル	1975年4月～現在	約36万件/年	理工学全般	JICST 発行の「科学技術文献速報」に対応
	CASEARCH 化学文献ファイル	1977年1月～現在	約45万件/年	化学・化学工業	米国CAS発行の“Chem. Abst.”に対応
	MEDLINE 医学文献ファイル	1974年1月～現在	約25万件/年	医学・薬学	米国国立医学図書館発行の“Index Medicus”に対応
	TOXLINE 毒性文献ファイル	1977年1月～現在	約12万件/年	毒物学とその関連分野	“Chem. Abst.” “Index Medicus” など8種に対応
	JICST 国内医学文献ファイル	1981年4月～現在	—	医学・薬学	「医学中央雑誌」などに対応
	BIOSIS 生物学文献ファイル	1979年1月～現在	約30万件/年	生物学	米国BIOSISの“Biological Abst.”などに対応
	CAB 農学文献ファイル	1979年1月～現在	約16万件/年	農 学	英国CAB発行の“Main Abst. J”など29種に対応
研 究 情 報 フ ァ イ ル	JICST 科学技術研究情報ファイル	1976年～現在	約2万件	理工学全般	日本国内の公共試験研究機関約540機関の研究テーマ
	SSIE 研究情報ファイル	1980年～	約10万件	理工学社会科学	米国内の約1,300機関の進行中、計画中の研究テーマ

(b) 検索補助用データベース(サポートファイル)

デ ー タ ベ ー ス 名	内 容
JICST 科学技術用語シソーラスファイル	JICST ファイルに使用されている統制語キーワードが収録されている。 JICST ファイルを検索するとき使用する。
JICST 資料所蔵目録ファイル	JICST ファイルを資料番号から検索するとき資料番号を入力する必要があるが、このファイルによって雑誌名からこの資料番号が分かる。
CASNAME 化合物ファイル	JICST が毎週発行されるCA SEARCHファイルから化合物名の巻末索引語部分だけを抜き取って作成するものである。したがって、このファイルの中化合物名はCASの正式名称だけから成り、同義語は含まれていない。
CHEMLINE 化合物ファイル	NLM(米国国立医学図書館)が作成するもので、主にTOXLINE ファイル中に含まれる化合物から成っている。化合物名のCAS正式名称だけでなく、同義語も含まれている。その他環情報なども含まれている。 TOXLINE, CA SEARCHファイルを検索するとき使用する。
MeSH 医学用語ファイル	NLMが作成するもので、MEDLINE ファイルに使用されている統制語キーワードが収録されている。MEDLINE ファイルを検索するとき使用する。

から成る質問を端末から登録・更新可能なようにしておくことができる。これにより、データベースの追加・変更時、自動的にバッチプログラムから呼ばれてSDI(Selective Dissemination of Information)検索も実行される。また従来、複写依頼は郵送により行なわれていたが、JOIS-IIでは端末から

表2 JOIS-II コマンド一覧表 JOIS-II には豊富なコマンドが用意されている。特に端末から、質問登録、複写受けが可能となり、会話機能のエンハンスとともにサービス機能も増した。

項 番	分 類	コマンド名称 (検索ステートメント含む)	有 無		概 要
			II	I	
1	検 索 機 能	完全一致検索ステートメント	○	○	入力タームによる完全一致検索
2	"	マスク文字検索ステートメント	○	○	入力タームによる前方一致検索、中間任意一致検索追加
3	"	論理式検索ステートメント	○	○	入力ターム、検索ステップ番号などによる論理式検索
4	"	制限式検索ステートメント	○	○	言語などによる回答件数の制限、ストリングサーチ機能追加
5	"	¥ OR	○	×	検索ステップ番号などの論理和簡易入力
6	"	¥ LINK	○	×	リンク付の論理式検索
7	検 索 補 助	¥ EXPAND	○	○	入力タームで前方一致するタグを出力、中間任意一致追加
8	"	¥ ECNT	○	○	先の¥ EXPANDの続きを出力 (旧¥ E99)
9	"	¥ REFER	○	○	保存されている検索ステップ、登録質問、登録サブヘディングの一覧表示
10	"	¥ CLEAR	○	○	保存されている検索ステップ、登録質問の消去
11	"	¥ LOOKUP	○	×	ディスクリプターなどをキーとしてJICSTシソーラスを参照
12	"	¥ MESH	○	×	MeSHタームをキーとして、MESHファイルからMESH Tree No.を参照
13	"	¥ SUBHEADING	○	×	サブヘディングの登録、削除
14	回 答 出 力	¥ DISPLAY	○	○	回答出力指示(24行以内)、ディスプレイ端末は無印字
15	"	¥ PRINT	○	○	回答出力指示
16	"	¥ OFFLINE	○	○	センターでの回答出力指示
17	会 話 機 能	¥ JOIS	○	○	JOISとの会話開始指示 (JOIS-I では¥ IN)
18	"	¥ END	○	×	JOISとの会話終了指示
19	"	¥ FILE	○	×	質問の開始指示(データベースを指示)
20	"	¥ QEND	○	○	質問の終了指示
21	"	¥ KEEP	○	×	質問を中断し一時的に退避させる
22	"	¥ RESTART	○	○	中断した質問の再開始指示
23	質 問 登 録	¥ SAVE	○	×	本コマンド投入以前の検索ステップを登録(後宣言形質問登録)
24	"	¥ STACK	○	×	1ステップ目からの検索ステップの登録指示(前宣言形質問登録)
25	"	¥ ENDSTACK	○	×	検索ステップの登録の終了
26	"	¥ EDIT	○	×	登録質問の修正
27	"	¥ QSAVE	○	×	単なる質問の登録を翌日以降保存
28	"	¥ CALL	○	×	登録質問の呼び出し実行
29	"	¥ TEST	○	×	SDI質問などのテストラン
30	会 話 補 助	¥ NEWS	○	×	ニュース、データベース案内などの表示
31	"	¥ COMMENT	○	×	端末利用者がシステム管理者に伝えるコメント入力
32	"	¥ USERS	○	×	同時接続中端末数を表示
33	"	¥ ORDER	○	×	複写申込み
34	端 末 管 理 者 用	\$ PASSWORD	○	○	ユーザーパスワードの登録・更新
35	"	\$ ACCOUNT	○	×	当月内(前日までの)料金累計の表示
36	"	\$ REFER	○	×	質問登録状況の表示

注：II (JOIS-II の意)、I (JOIS-I の意)、SDI の意 図2略語説明参照

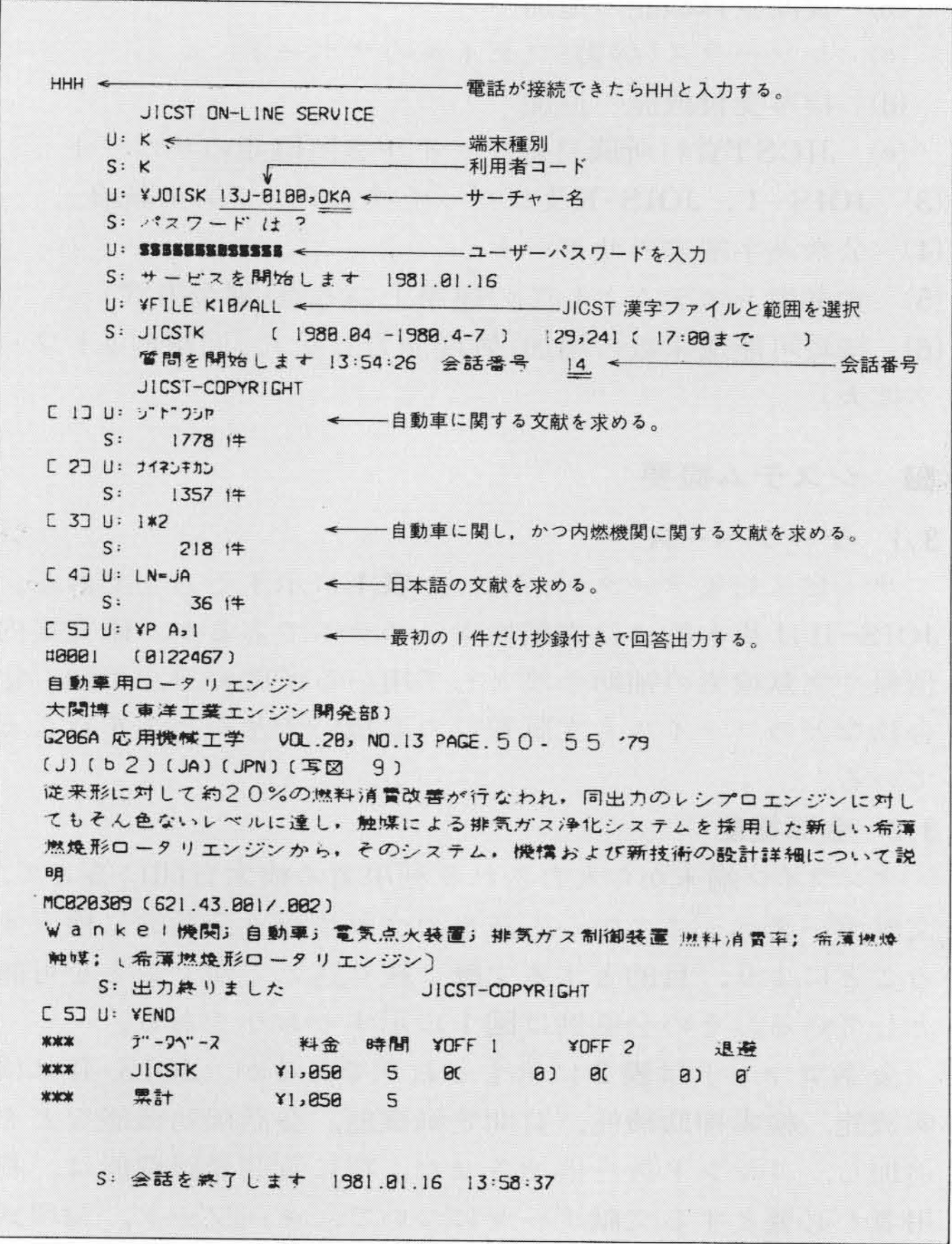


図1 JOIS-IIの会話例(JICST漢字ファイル検索例) 公衆漢字端末から「自転車について書かれていて、かつ内燃機関について書かれている文献」を検索した場合の例を示す。

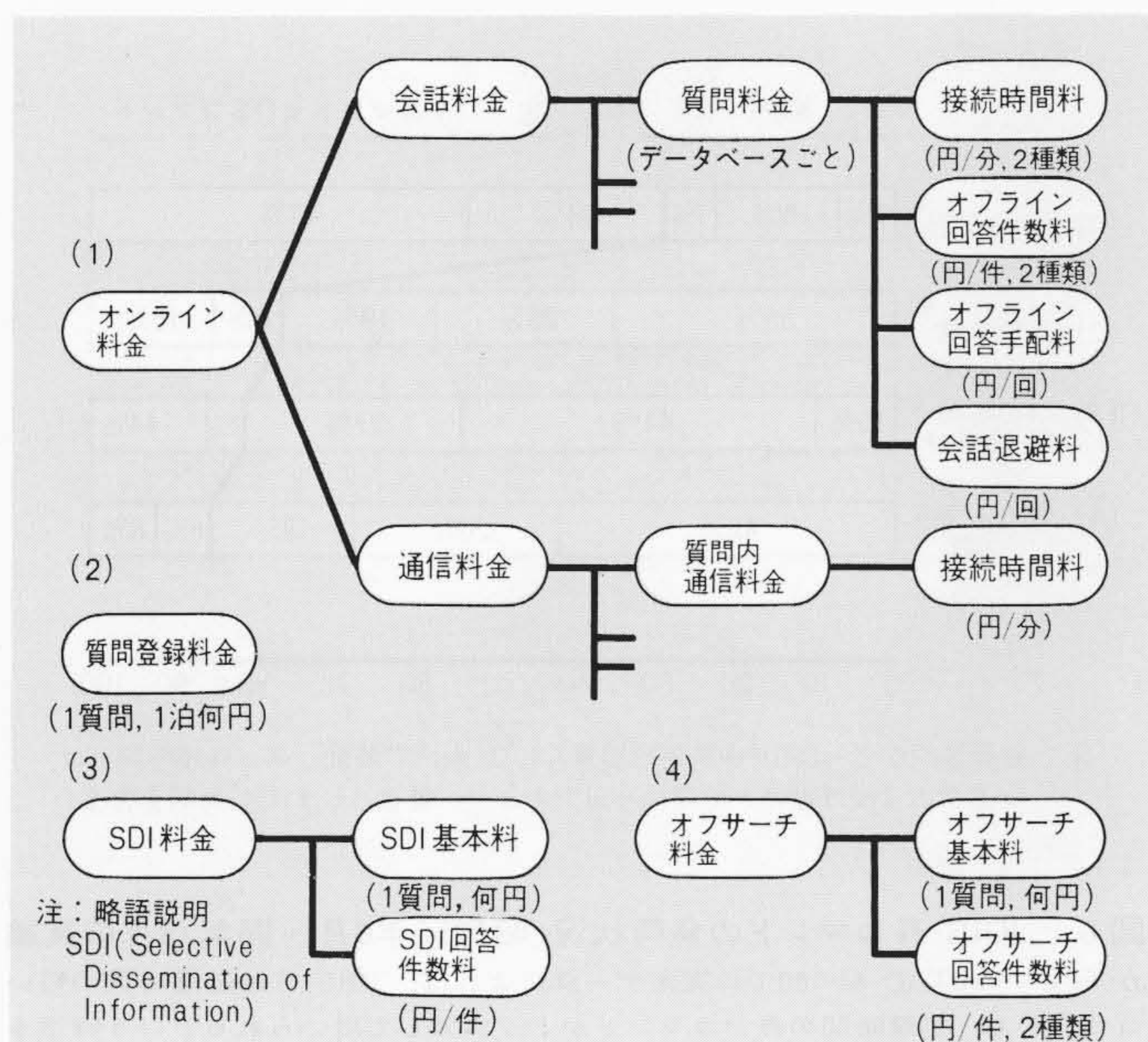


図2 JOIS-II 課金体系 オンラインの課金基準は簡明さを特徴としており、端末接続時間を基準とした従量制の課金基準を採用している。

申し込むことができるようになり、一貫した情報サービスが行なえるようになった。

3.3 課 金

JOIS-IIの課金体系は、図2に示すとおりである。オンラインの課金基準は簡明さを旨とし、端末接続時間を基準とした従量制の課金基準を採用している。質問登録料金は1質問、1泊何円という形で課金される。また、将来、DDX(Digital Data Exchange)網と接続されたときには、その分の課金も配慮してある。

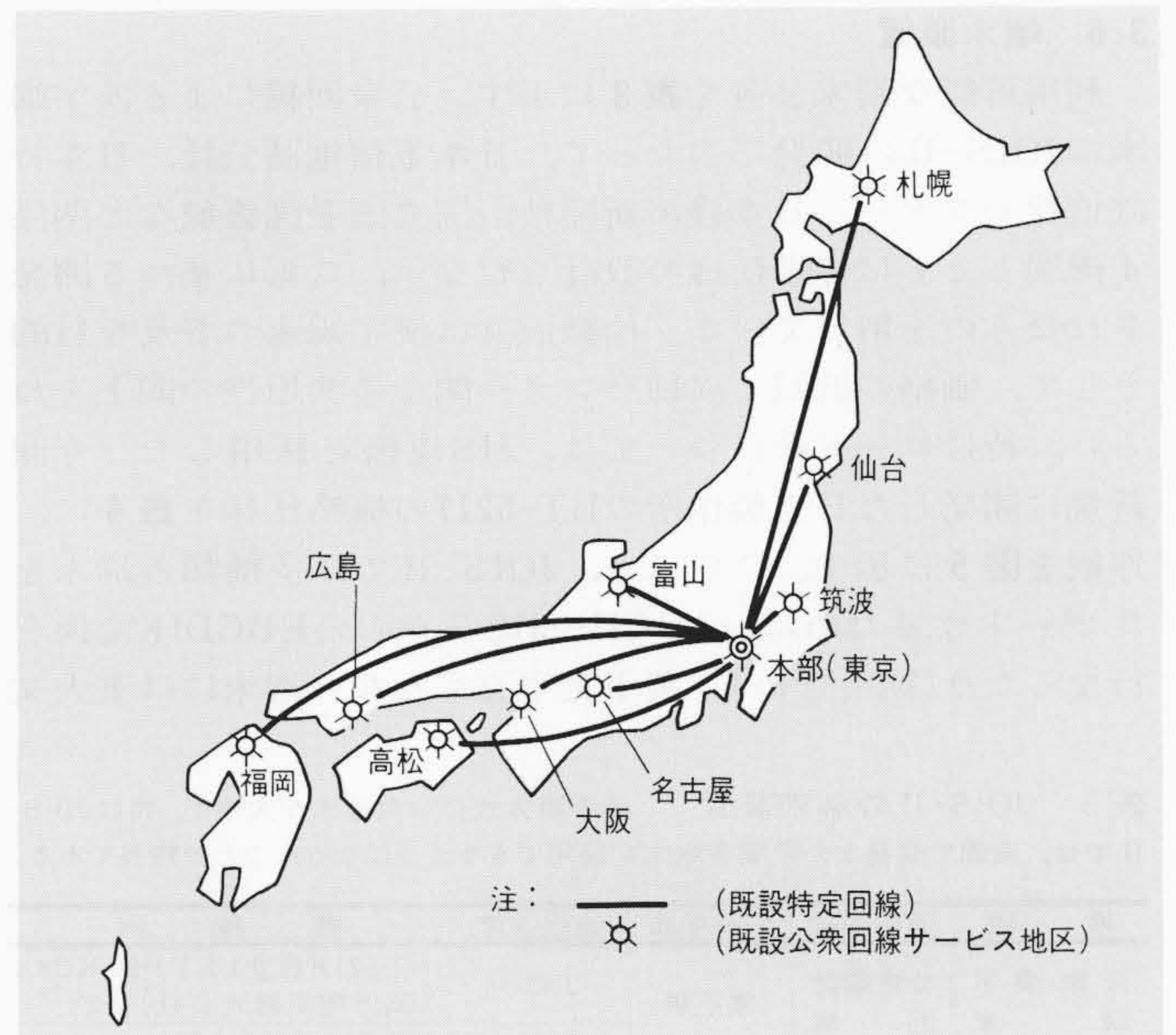


図4 JOIS-II 通信網 東京にセンターを置き、全国10箇所に拠点をもつ。センターと遠隔地の間は高速の特定回線で結ばれ、多重集配信装置により時分割多重利用している。

3.4 JOIS-IIソフトウェア構成

オンライン部のソフトウェア構成を図3に示す。

3.5 通信網

JOIS-IIの通信網は図4に示すとおりになっている。通信網の各ノードには、公衆回線を集信する集信装置を設置している。中央とノードの間は特定通信回線D-1回線で結び、通信速度9,600bpsなどで利用し、多重集配信装置により時分割多重利用している。また、将来はDDXのパケット交換網を利用したサービスも予定している。

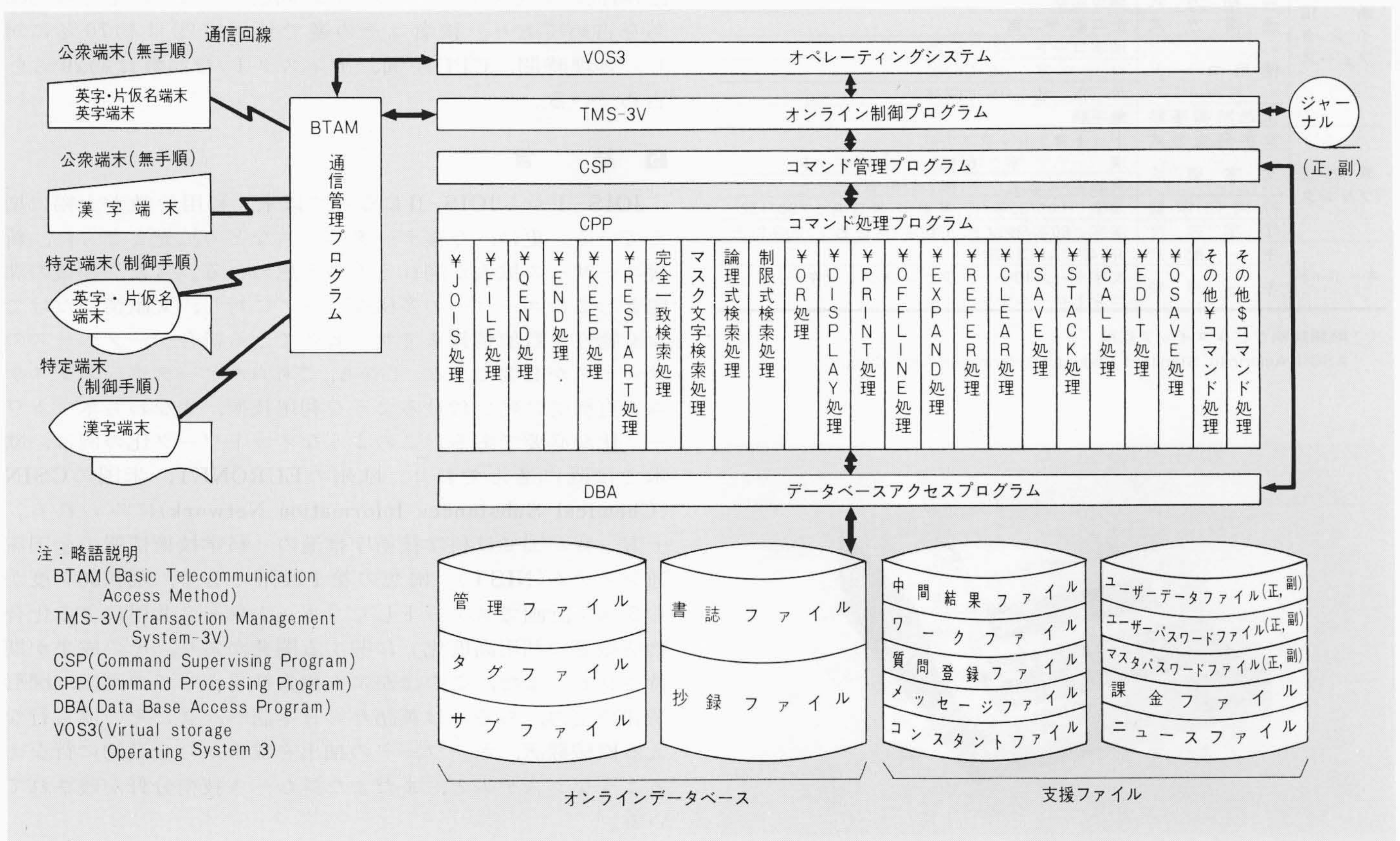


図3 JOIS-IIソフトウェア構成 ソフトウェア構成はVOS3, BTAM, TMS-3V, MPPから成り、MPPは更にCSP, CPP, DBAから成る。

3.6 端末装置

利用可能な端末装置を表3に示す。公衆回線による漢字端末はJOIS-IIの開発に当たって、日本電信電話公社、日本特許情報センター、日本経済新聞社、国立国会図書館など関係4機関とともに標準仕様の設計を行ない、これに基づき開発されたものを用いている。仕様設計は漢字端末の普及を目的として、価格の低減と異種センター間での共用性の向上をねらい、特にコード系については、JIS規格を採用した。今回新規に開発した日立製作所のHT-5217の概略仕様を表4に、外観を図5に示す。このほか、JOIS-IIでは多種類の端末をサポートするために、ASCIIやJIS 7からのEBCDIK変換を行ったり(逆も含め)、英小文字をもたない端末には英大文

表3 JOIS-IIの端末装置 無手順方式の公衆端末が大半で、特にJOIS-IIでは、安価で簡易な公衆漢字端末が使用できるようになったことが特長である。

種類	通信回線	制御手順	通信速度	機種例
公衆漢字端末	公衆電話回線	無手順	300bps	HT-5217(日立), STT-20IKCK(三洋電気株式会社)など
			1,200bps*	HT-5217(日立)など
公衆英字・片仮名端末	公衆電話回線	無手順	300bps	Silent745(米国テキサスインスツルメント社)など
			1,200bps*	HT-5217(日立)など
特定漢字端末	特定通信回線	HSC3	2,400bps	HT-590/30(日立)など
特定英字・片仮名端末	特定通信回線	HSC3	2,400bps	H-9415(日立)など

注：*1,200bpsは、未サービスである。

表4 HT-5217漢字プリンターミナル仕様一覧 300bpsだけでなく1,200bpsもサポートしているのが最大の特徴である。このほかにも液晶ディスプレイによる入力データの表示、行単位印字出力などの特徴をもつ。

項目	機能	仕様
液晶ディスプレイ	表示文字数	40字×2行 1行目…入力データ表示(100字シフト表示) 2行目…状態表示
通信インタフェース	データ通信速度	300bps, 1,200bps
	同期方式	調歩同期
	通信方式	全二重/半二重
	使用コード	漢字コード：JIS C6226
		ローマ字：JIS C6220
		片仮名：JIS C6220
感熱式プリンタ	伝送制御手順	無手順
	文字発生方式	ドットマトリックス方式
	文字構成	漢字：16(横)×18(縦)ドット 英数字・片仮名：7(横)×18(縦)ドット
	行内印字数	漢字：*40/66字/行, 英数字・片仮名：*80/132
	印字速度	漢字：60字/秒以上, 英数字・片仮名：120字以上
キーボード	キー配列	JIS4段/ASCII2段
	キー手順	文字キー, 10キー, ファンクションキー, コントロールキー

注：略語説明など*スイッチ切換
ASCII(American Standard Code for Information Interchange)



図5 HT-5217漢字プリンターミナルの外観 無制御手順で接続可能な小形、かつ軽量のターミナルである。

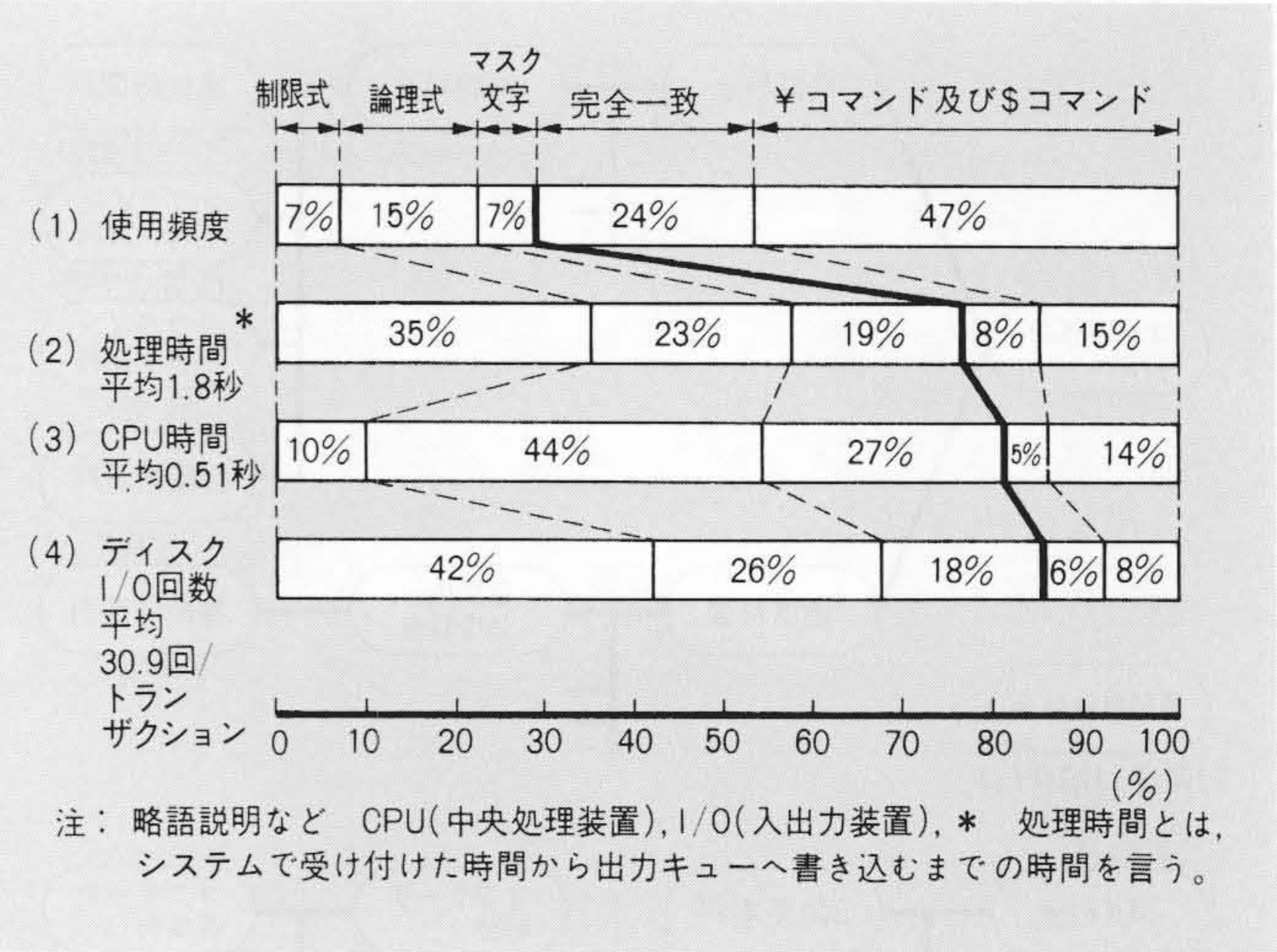


図6 JOIS-II コマンドの負荷状況(昭和56年6月～同年/8月の実績から) HITAC M-180での実測データによれば、JOIS-IIは処理時間の短いコマンドと、処理時間の長いコマンドがミックスして用いられるという特徴をもっている。

字にして出力することができる。更にメッセージも漢字、英語、片仮名と3種用意され、操作性も配慮されている。

4 稼動評価

昭和56年6月から8月までのHITAC M-180による実測結果をもとに作成したコマンド負荷状況を図6に示す。JOIS-II コマンドは制限式、論理式、マスク文字コマンドにみられるような処理時間の長いコマンド群と、反対に完全一致などにみられる処理時間の短いコマンド群がミックスして用いられるという特徴をもっている。前者のコマンドの使用頻度は合わせて約30%にもかかわらず、処理時間、CPU(中央処理装置)時間、ディスクI/O(入出力)回数はそれぞれ全体の約80%を占めており、後者はその逆で使用頻度は約70%に対し、処理時間、CPU時間、ディスクI/O回数は約20%を占めている。

5 結 言

JOIS-I からJOIS-II になって以来、利用件数が大幅に増えている。更に、今後データベースなどの拡充とともに、新規ユーザーの拡大が図れるものと思われる。また、今後の課題としてはユーザーの多様なニーズに対し、文献情報だけでなく物性・数値情報までサービスできる総合データベースのサービスが必要である。しかも、これらのデータベースシステムを有機的に結び付けるような利用技術、すなわちネットワーク化が必要である。このようなネットワーク化の例は、欧米では既に進んでおり、欧州のEURONET、米国のCSIN (Chemical Substances Information Network)にみられる。一方、我が国では科学技術庁推進の「科学技術情報の全国流通システム(NIST)」構想の第1段階として、昭和56年度から5箇年計画でスタートした「ネットワーク共用による化合物情報等の利用高度化」に関する開発があり、その成果が期待される。また、このほかにも検索結果としての画像、図形情報の出力、あるいは英語から日本語へ、またその逆も行なえる機械翻訳、キーワードの抽出を端末から会話的に行なえるようなシステムと、まだまだ挑むべき技術分野が残されている。

終わりに、このシステム建設に当たっていろいろ御指導、ご協力をいただいた関係各位に対して深く感謝する次第である。