International Journal of Thermophysics, Vol. 4, No. 4, 1983

Development of a Database System for the Thermophysical Properties of 340 Pure Fluids

T. Makita,[1] Y. Tanaka,[1] T. Takamori,[2] and Y. Takaoka[3]

Received December 16, 1983

A new database system was designed and constructed in order to provide the newest and most reliable numerical data on thermophysical properties of fluids to the scientific and technical community as quickly as possible. The database is concerned with 12 kinds of property data: density, specific volume, compressibility factor, vapor pressure, specific heat, specific enthalpy, specific entropy, viscosity, thermal conductivity, surface tension, refractive index, and dielectric constant of 340 pure fluids under all conditions of seven convenient physical states. Only the most probable data, which were critically evaluated in advance, are compiled as functions of temperature and pressure in the form of either numerical data tables or correlation formulas. Property values at desired grid points of temperature and pressure can be obtained in SI units directly or by interpolations with some supplementary information, such as the grade of reliability, source bibliography, and fundamental physical constants of the fluid. The database is usable as conversational processing by TSS, the tabulation of property data in a batch processing, and one of the subroutine libraries in the computation of a user's program. Ten kinds of application programs are prepared for multipurpose retrievals so that anyone can use the database without any special knowledge of the structure or machine languages of this system.

KEY WORDS: compressibility factor; database system: density; dielectric constant; refractive index; retrieval program; specific enthalpy; specific entropy; specific heat capacity; specific volume; surface tension; thermal conductivity; vapor pressure; viscosity.

1. INTRODUCTION

Accurate information on the thermophysical properties of fluids is essential both to the research work of scientists and to the development of industrial

[1]Department of Chemical Engineering, Kobe University, Kobe 657, Japan.
[2]Department of Instrumentation Engineering, Kobe University, Kobe 657, Japan.
[3]Information Processing Center, Kobe University, Kobe 657, Japan.

0195-928X/83/1200-0283$03.00/0

technologies. Although an enormous amount of property data is generated and published every year, there often exist unexpected large discrepancies among the experimental data, reported by different investigators, of a substance under the same conditions. Therefore it is quite risky to use raw data, which one happened to discover, without critical evaluation. However, the recent "information explosion" has made it quite difficult and troublesome for an individual scientist to extract the necessary data completely from the literature and assess the quality of the published information. Thus, the collection, critical evaluation, organization, and dissemination of numerical data are indispensable together with processing of the published literature, in order to improve scientific communication. Various data centers or information analysis centers are working on the evaluation and synthesis of numerical data, in order to provide the best values to the community of science and technology. However, the traditional mechanisms for scientific communication, such as academic journals, abstracting and indexing services, and so on, have all demonstrated a cultural lag in accommodating to the present requirements.

In consideration of this situation, a new database system on the thermophysical properties of fluids has been designed and constructed. The primary object of the database is to provide reliable selected values of thermophysical properties of fluids, which have been critically evaluated in advance, to the scientific and technical community as accurately and quickly as possible. The present database is concerned with 12 kinds of thermophysical properties: density, specific volume, compressibility factor, vapor pressure, specific heat, specific enthalpy, specific entropy, viscosity, thermal conductivity, surface tension, refractive index, and dielectric constant of 340 important fluids. The basic philosophy of the database and the principle of its construction were reported previously [1,2]. In the present paper, the general features and functions of this database system, including various retrieval methods, are described in detail.

2. FEATURES OF THE PRESENT DATABASE

The distinguishing features in the construction of the database are summarized as follows.

1. Only the most probable values, which were critically evaluated in advance, have been compiled in the database.
2. The fluids registered were selected on reflection of practical demands, including 49 elements and 111 inorganic and 180 organic compounds. According to the demands in the future and the availability of selected data, new substances will be added at any time.

Table I. Codes of Thermophysical Properties

Code	Property	Unit
D	Density	$kg \cdot m^{-3}$
V	Specific volume	$m^3 \cdot kg^{-1}$
Z	Compressibility factor	
P	Vapor pressure	bar(= 0.1MPa)
C	Specific heat	$kJ \cdot kg^{-1} \cdot K^{-1}$
H	Specific enthalpy	$kJ \cdot kg^{-1}$
S	Specific entropy	$kJ \cdot kg^{-1}K^{-1}$
R	Viscosity	$Pa \cdot s$
K	Thermal conductivity	$W \cdot m^{-1} \cdot K^{-1}$
G	Surface tension	$N \cdot m^{-1}$
N	Refractive index	
E	Dielectric constant	

3. Twelve kinds of thermophysical properties listed in Table I were assigned from the viewpoint of scientific and technical requirements.
4. All the possible physical states, where a substance exists as gas or liquid, were defined for user's convenience as listed in Table II.
5. The property data are compiled as a function of temperature in K and pressure in bar in the form of either numerical data tables or correlation formulas. The property data were entered in their original units and converted into SI units in this database system.
6. Each set of property data is accompanied with the data source bibliography and the grade of estimated reliability, whose codes are given in Table III. The uncertainties of the data were determined by critical evaluation, taking account of the inaccuracy of the original work and judging the possible departures from auxiliary data taken from other sources.

Table II. Codes of Physical States

Code	State
GO	Ideal gas
GA	Gas at 1.01325×10^5Pa
LA	Liquid at 1.01325×10^5Pa
GS	Gas at saturation
LS	Liquid at saturation
GP	Gas at high pressures
LP	Liquid at high pressures

Table III. Codes of Data Grade

Code	Range of uncertainty
0	Unclassified
1	Below 0.1%
2	0.1–0.5%
3	0.5–1.0%
4	1.0–2.0%
5	2.0–3.0%
6	3.0–5.0%
7	5.0–10.0%
8	Above 10.0%

7. Each fluid is accompanied with the fundamental physical constants, such as molecular weight, normal melting and boiling points, critical temperature, pressure, and volume and density as "additional information."

3. STRUCTURE OF THE DATABASE SYSTEM

3.1. Database Management System (DBMS)

The present database system has been constructed by means of the database management system "Information Query" (INQ), which is one of the specific database techniques of the computer ACOS Series 77 NEAC System 1000 produced by Nippon Electric Co. Ltd. at the Information Processing Center of Kobe University. INQ [3, 4] is an inverted file type DBMS, and its special features are as follows:

1. INQ is one of DBMS which belongs to a relational type. Therefore it is easy to design the database system because INQ is able to represent naturally the relations among compiled data in the form of tables.
2. It is easy to add new data or to revise the compiled information by the partial revisions of the INQ file concerned, because individual INQ files are quite independent of each other. The addition or change of the retrieval routines is also possible by the revision of an INQ section which combines the INQ files.
3. It is an important merit that INQ has the FORTRAN language as the host language, because data processing is indispensable in the retrieval routines of the present database system.

3.2. Structure of the Database

The database consists of four INQ files as shown in Fig. 1. Each file is combined mutually with substance codes, reference codes, or formula codes. As shown in Table IV, the data structures in INQ files are described

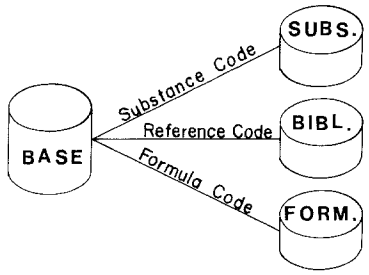

Fig. 1. Composition of the thermophysical property database. BASE: Base file (property data information file in Table IV). SUBS: Substance information file. BIBL: Bibliography information file. FORM: Formula type information file.

in a "file description language" (FDL) which is similar to the COBOL language in data description. The data attributes X, 9, and FB denote characters, figures (unpacked decimal), and a floating binary, respectively. PIC denotes picture. In the data type, PKY and DSP denote a primary key

Table IV. FDL Description of INQ Files

Level number	Data item	Data attribute	Data type
	Property data information file		
02	Substance code	PIC X(4)	
02	Property code	PIC X(1)	
02	Physical state code	PIC X(2)	
02	Reference code	PIC X(5)	
02	Grade code	PIC 9(1)	DSP
02	Property data tables (N)		
03	Pressure	PIC FB	DSP
03	Data sets in a table (N)		
04	Temperature	PIC FB	DSP
04	Property data	PIC FB	DSP
02	Formula information (N)		
03	Lower limit of T	PIC FB	DSP
03	Upper limit of T	PIC FB	DSP
03	Lower limit of P	PIC FB	DSP
03	Upper limit of P	PIC FB	DSP
03	Formula code	PIC 9(2)	DSP
03	Minimum power on T	PIC 9(2)	DSP
03	Maximum power on T	PIC 9(2)	DSP
03	Minimum power on P	PIC 9(2)	DSP
03	Maximum power on P	PIC 9(2)	DSP
03	Applying condition of formula	PIC 9(1)	DSP
03	Coefficients of formula (N)		
04	Values of coefficients	PIC FB	DSP

Table IV. Continued.

Level number	Data item	Data attribute		Data type
	Substance information file			
02	Substance code	PIC	X(4)	PKY
02	Chemical formula thesaurus (N)			
03	Chemical formula	PIC	X(20)	
02	Representative substance name	PIC	X(40)	DSP
02	Substance name thesaurus (N)			
03	Substance name	PIC	X(40)	
02	Chemical abstract registry number	PIC	X(20)	
02	Molecular weight	PIC	X(7)	DSP
02	Melting point	PIC	X(9)	DSP
02	Boiling point	PIC	X(9)	DSP
02	Critical pressure	PIC	X(8)	DSP
02	Critical temperature	PIC	X(8)	DSP
02	Critical volume	PIC	X(8)	DSP
02	Critical density	PIC	X(7)	DSP
	Bibliography information file			
02	Reference code	PIC	X(5)	PKY
02	First author	PIC	X(20)	
02	Second author	PIC	X(20)	
02	Third author	PIC	X(20)	
02	Fourth author	PIC	X(20)	
02	Title	PIC	X(160)	DSP
02	Journal or publication	PIC	X(100)	DSP
02	Year	PIC	9(4)	DSP
	Formula type information file			
02	Formula code	PIC	9(2)	PKY
02	Formula	PIC	X(130)	DSP

and a display item. The items which are not assigned to DSP constitute the retrieval items. These items are usable as a retrieval key since the inverted index is prepared for them. In order to minimize the load of DBMS in data loading processes, the retrieval items are limited to essential ones. In the property data information file, a set of property data in table form or formula type constitutes one record. Although each record has two data items, that is, "property data tables" and "formula information," the property data are compiled actually in either of them. The mark (N) in the data item denotes an indefinite repeating group.

In order to retrieve data in the present database system, it is required to construct the "INQ sections" which supply a suitable user view by combining several INQ files. The present retrieval system has seven INQ sections, some of which are selected properly according to a retrieval purpose.

4. RETRIEVAL SYSTEM

4.1. Usage of the Database

The primary object of the present database is to provide reliable selected values of thermophysical properties of fluids to the scientific and technical community as quickly and accurately as possible. Therefore, as it is clearly not sufficient to store a large quantity of collected information in the database, it is rather important to retrieve the desired information from the database effectively and to disseminate it quickly. In order to fulfil this purpose, an orginal retrieval system joined to the present database has been designed and prepared for users' convenience. The configuration for the utilization of this database system is shown in Fig. 2.

The retrieval searches are made not only through TSS (time sharing system) terminals but also by batch processings with line-printers. The retrieval routines through a batch processing or a remote batch processing are also available, in order to use this database as one of the subroutine libraries in the computation of user's programs.

4.2. Retrieval Routines

The present retrieval system consists of 10 retrieval routines which are considered to be helpful in common usage, as listed in Table V. The retrieval routines are classified into three categories from their specific functions.

4.2.1. Retrieval Routines for Registered Information

Before the retrieval of desired information it is necessary to know first (i) whether the desired information is compiled in the database, and (ii) how the information is stored in the database, if available. Since the compiled information is not fixed permanently, it is desirable to retrieve this routine first every time. The present database system has four retrieval

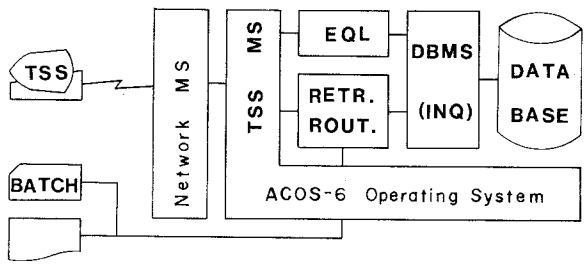

Fig. 2. Configuration for utilization of the database system. MS: Management system. EQL: End user's language of INQ. RETR. ROUT.: Retrieval routines.

Table V. Retrieval Routines by TSS[a]

Code	Function	Input Information	Output Information
*R*EGI	Search for a REGIstered fluid	Fluid code/chem. formula/CA reg. no./ English names	Fluid code, chem. formula, CA reg. no., English names.
*C*OMB	COMBination table of properties vs physical states, where data are available	Fluid code	Fluid code, formula, name, combination table of properties vs. physical states, where available status is shown by asterisk(*)
*D*ATA	Search for registered DATA-information for a property of a designated fluid	Fluid code, Property code/ "All"	Fluid code, formula, name, property code, phys.-state code, ranges of registered data, form of data, grade code, source ref. no., year
*I*NDI	Search for registered information for an INDIvidual property of all fluids	Property code, Phys.-state code/ "All"	Property code, phys.-state code, fluid codes, chem. formulas
*A*DDI	Output of ADDItional information (fundamental physical constants) of a designated fluid	Fluid code	Fluid code, formula, name, mol. wt., melting point, boiling point, critical values
*L*ITE	Output of source LITErature information	Ref. no.	Ref. no., bibliography (authors, title, journal, vol., pages, year), fluid codes and property codes involved
*B*IBL	Output of BIBLiographies from author's name	Author's name	Author's name, ref. nos., bibliographies
*P*ROP	Retrieve a PROPerty value at a designated condition	Fluid code, property code, phys.-state code, temperature, pressure	Retrieval condition, property value, phys.-state, error status, information of used data in the interpolation processing
*T*ABL	Output of a property data TABLe for a designated fluid at given phys.-state	Fluid code, property code, phys.-state code, temperature range, its increment, pressure range, its increment	Fluid code, formula, name, additional information, property code, phys.-state code, grade, ref. no., property data table, bibliography

Table V. Continued.

Code	Function	Input Information	Output Information
*S*UBS	Retrieve SUBStances which have the designated property values at a given condition	Property code, phys.-state code, temp., press., range of the property values, no. of retrieval fluids/"All"	Retrieval condition, fluid code, formula, name, property value, phys.-state

[a] In "Code" column, the first letter written in italic is effective in the retrieval. In "Input Information" column, slash "/" means "or".

routines REGI, COMB, DATA, and INDI for registered information such as substance, fluid code, chemical formula, available range of property data, form of input data, and so on.

4.2.2. *Supplementary Retrieval Routines*

The database has three supplementary retrieval routines, ADDI, LITE, and BIBL for the user's convenience. LITE and BIBL are concerned with the data source bibliography such as authors, title, journal, volume, page, and year. The present database also compiles "additional information" on the fundamental physical constants, such as molecular weight, normal melting and boiling points, critical temperature, pressure, volume, and density for each substance.

4.2.3. *Main Retrieval Routines*

The database is provided with three main retrieval routines: PROP and TABL for the retrieval of property data, and SUBS for the retrieval of substances which have designated property values at an assigned condition. When a set of substance, property, physical state, temperature, and pressure is designated, a property value is given through PROP directly from the compiled data tables or by interpolations, together with the status information on used data, method of interpolation, and error in data processing. This routine is also available in a batch processing as one of the FORTRAN subroutines in users' arbitrary programs.

TABL is concerned with the output of property data tables. If a set of substance, property, physical state, temperature and pressure ranges, and their intervals in the table are designated, data tables are obtained, together with the additional information of data source bibliography and formula information, if the data are calculated from correlation formula. When the property data were compiled in the form of tables, the compiled data are obtained without any data processing in order to maintain the high retrieval efficiency without decreasing their accuracy.

The retrieval routine SUBS is concerned with the retrieval of substances whose property values are within a designated range at a designated condition of temperature and pressure. When the property, the range of property values, physical state, and temperature and pressure are designated, all the corresponding substances are obtained with the property values.

4.3. Method of Data Processing

The interpolation of property data is one of the main data processings in the retrieval system. When the property data were compiled in the form of correlation formulas, it is performed easily by the formula processing routine. This routine calculates the designated property values directly within the effective range of the formula. If the property data were compiled in the form of tables, it is often necessary to interpolate the property value at a designated condition using the available data. The interpolation of data is performed generally by the Lagrange method of interpolation using four data points. When the property value changes remarkably with temperature or pressure, the logarithmic property values are used for the interpolation. Each calculated value thus obtained is accompanied with some status information concerning the data processings. Estimation of property values by extrapolation outside the registered range of compiled data tables or the effective range of correlation formulas is forbidden in the present database, from the viewpoint of accuracy of the data.

4.4. Examples of Retrieval

Some examples of the retrievals through a TSS terminal are shown in Figs. 3, 4, and 5. If one wants information on methane, for example, the substance code is 1001, and one retrieves first by COMB, as shown in Fig. 3.

```
    -- T H E R M O P H Y S I C A L   P R O P E R T I E S --          11/21/83
  ***** REGISTERED PROPERTY-PHYSICAL STATE COMBINATION TABLE *****
SUBSTANCE CODE ? 1001

                                          PROPERTY DATA INFORMATIONS
 S U B S T A N C E                      ( * : PROP. DATA ARE REGISTERED.)
                                        ( - : COMBINATION IS MEANINGLESS.)
                                            D V Z P C H S R K G N E
 CODE:    1001                          GO - - - - * * * - - - - - -
                                        GA * * * - * * * * * -
 FORMULA: CH4                           LA * * * - * * *
                                        GS * *   * * * * * * -
 NAME:    METHANE                       LS * *   * * * * * * *
                                        GP * * * - * * * * * -
                                        LP * * * - * * * * *
```

Fig. 3. Example for retrieval of available property data by COMB.

```
      -- T H E R M O P H Y S I C A L   P R O P E R T I E S --          11/21/83

   ***** PROPERTY DATA RETRIEVAL *****

SUBSTANCE CODE ? 1001
PROPERTY CODE ? R
PHYSICAL STATE CODE ? GP
TEMPERATURE (IN K) ? 348
PRESSURE (IN BAR) ? 20

** RETRIEVAL CONDITIONS **

  SUBSTANCE CODE       :  1001
            FORMULA    :  CH4
            NAME       :  METHANE
  PROPERTY CODE        :  R ( VISCOSITY )
  PHYSICAL STATE CODE  :  GP ( GAS UNDER HIGH PRESSURE )
  TEMPERATURE          :  348.00 ( K )
  PRESSURE             :   20.00 ( BAR )

** RESULTS OF RETRIEVAL **

  PROPERTY DATA ( VISCOSITY )              :  0.130320E-04 ( PA.S )
**
  PHYSICAL STATE OF THE SUBSTANCE          :  GAS
  ERROR STATUS :   0 ( NO ERROR IN CALCULATION )

** USED DATA **

  DATA TYPE                                :  TABLE
  PHYSICAL STATE CODE OF DATA              :  GP ( GAS UNDER HIGH PRESSURE )
  GRADE OF DATA                            :  5  ( 1 -- 2 % IN ERROR )
  SOURCE CODE OF DATA                      :  00503
  NO. OF DATA POINTS USED FOR INTERPOL.:  2 POINTS
  TEMP. INTERVAL BETWEEN 2 DATA POINTS :  10.00 ( K )
  PRES. INTERVAL BETWEEN 2 DATA POINTS :   0.   ( BAR )
  LOGARITHMIC INTERPOLATION                :  UNUSED
```

Fig. 4. Example for retrieval of a property value at a designated condition by PROP.

```
      -- T H E R M O P H Y S I C A L   P R O P E R T I E S --          11/21/83

   ***** SUBSTANCE RETRIEVAL *****

PROPERTY CODE ? R
PHYSICAL STATE CODE ? GP
TEMPERATURE (IN K) ? 348
PRESSURE (IN BAR) ? 20
PROPERTY DATA RANGE (IN SI) ? 0.000012,0.000015
NUMBER OF RETRIEVING SUBSTANCES ? ALL

** RETRIEVAL CONDITIONS **

   PROPERTY:         R ( VISCOSITY )
   PHYSICAL STATE:   GP ( GAS UNDER HIGH PRESSURE )
   TEMPERATURE:      348.00 ( K )
   PRESSURE:          20.00 ( BAR )
   PROPERTY DATA:    0.120000E-04  --  0.150000E-04 ( PA.S )
   NO. OF RETR. SUBS.:  ALL

** RETRIEVED SUBSTANCES **

        S U B S T A N C E                                  PROP. DATA      PHYS.
**
   CODE   FORMULA        N A M E                           R ( PA.S )      STATE
   1001   CH4            METHANE                           0.130320E-04    GP
   1003   C2H4           ETHYLENE                          0.123400E-04    GP
   1002   C2H2           ACETYLENE                         0.126970E-04    GP
```

Fig. 5. Example for retrieval of substances which have the designated property values at a condition by SUBS.

The database provides a combination table of properties versus physical states, where an asterisk shows that data are available. The user then retrieves the viscosity of gaseous methane at 348 K and 20 bar by PROP, as shown in Fig. 4. The database gives the viscosity value 0.13032×10^{-4} Pa · s for this retrieval. In Fig. 5 the user retrieves the substitutes for gaseous methane which have viscosity values between 0.12×10^{-4} and 0.15×10^{-4} Pa · s at 348 K and 20 bar by SUBS. Then he finds two substitutes, ethylene and acetylene, as well as methane with their viscosity values at the designated condition.

5. SUMMARY

A new database system was designed and constructed, through which one can retrieve thermophysical properties of fluids quite easily without any special knowledge of the structure or machine languages of the system. Since 1981 the database has been opened to the public tentatively, in order to examine the system and retrieval routines. However, taking into consideration the extensive demands of users, the present situation is not sufficient in the numbers of substances and properties adopted. Although the selection and evaluation of input data are difficult and time-consuming, the addition of new data will be continued, as well as the maintenance of the database, in order to keep the compiled data up-to-date. Furthermore, as the present database deals only with the properties of pure substances, the authors are now studying expansion of the present database to include fluid mixtures, which are important in industrial applications.

ACKNOWLEDGMENTS

This investigation has been partly supported by a Scientific Research Grant of the Ministry of Education, the Research and Development Funds of the Hyogo Foundation for Promotion of Science and Technology, and the Information Processing Center, Kobe University. The authors wish to thank Mr. M. Tanaka and Miss K. Hamaoka at Kobe University for their extensive work in the data processing.

REFERENCES

1. T. Makita, Y. Tanaka, T. Takamori, and Y. Takaoka, in *Data for Science and Technology*, P. S. Glaeser, ed. Proc. 7th Int. CODATA Conf., Kyoto, Japan, 1980. Pergamon Press, Elmsford, N.Y., 1981.
2. Y. Takaoka, T. Takamori, T. Makita and Y. Tanaka, *Trans. Inform. Process. Soc. Japan* **23**:406 (1982), in Japanese.
3. NEC Corporation, INQ General Description (FFK01E) (1980).
4. M. Hashimoto, T. Goto, K. Takeuchi, S. Mabuchi, and T. Doi, Database Management System "INQ (Information Query)", *NEC Res. Dev.*, No. 58 (1980).