



United Nations
Environment Programme



Food and Agriculture Organization
of the United Nations

Distr.
GENERAL

UNEP/FAO/PIC/ICRC.3/INF.4
11 January 2002

ENGLISH ONLY

INTERIM CHEMICAL REVIEW COMMITTEE

Third session

Geneva, 17 – 21 February 2002

Item 6 of the provisional agenda *

INFORMATION NOTE ON CHEMICAL ABSTRACT SERVICE (CAS) NUMBERS

Note by the Chair of the Interim Chemical Review Committee (ICRC)

1. The Interim Chemical Review Committee, at its second session in Rome recommended that that Chemical Abstract Service (CAS) numbers and precise chemical descriptions be used in identifying chemicals subject to the Interim PIC procedure. The Chair of the Committee also suggested that a presentation on the subject of CAS numbers be made at the third session of the Committee.
2. The Intergovernmental Negotiating Committee (INC), at its eighth session, considered issues arising out of the second session of the Interim Chemical Review Committee (UNEP/FAO/PIC/INC.8/7). This included the recommendation on the use of Chemical Abstract Service (CAS) numbers and precise chemical descriptions in identifying chemicals subject to the interim PIC procedure. The Committee concluded its consideration of this topic with a request that Governments, when submitting notifications of final regulatory action, describe all chemicals accurately by chemical name and CAS number
3. Annexed to the present note is general information on Chemical Abstract Service (CAS) numbers compiled by the Chair of the ICRC. This includes general background information and an overview on several Chemical Registries numbering systems as well as a general overview of the Chemical Abstract Service (CAS) system, taken from CAS Internet web site.

* UNEP/FAO/PIC/ICRC3/1

Appendix I
Chemical Registries - the Chemical Abstract Service (CAS)
original Author: Robert E. Buntrock
excerpt prepared by the Chair of the ICRC, R. Arndt

Background:

Chemical information has two unique features compared to all other species of information:

- 1) chemical structure, and
- 2) chemical reactions, which include chemical structures, physicochemical entities, and vector functions.

However, chemical structure is not always well defined and even if it is, precise nomenclature is required to define chemical entities. It soon becomes obvious that neither structure nor precise chemical nomenclature is that utilitarian, especially for use by the non-technical (or non-chemical) public, but even for effective, precise communication between chemists. Examples of nomenclature difficulties in the public sphere abound, including a very prevalent confusion of "silicon" and "silicone", a problem confounded by different spellings for either entity in some languages other than English.

For decades, chemists have responded by developing lists of chemical compounds, with associated attributes. Such lists are chemical registries and the individual compounds can be effectively described by an identity – ID - number. Although commercial chemical catalogues still list their "registries" by nomenclature – with all of the associated problems – compounds are then ordered by the catalogue number, making it a registry number.

Organisations in the business of making new chemical compounds, especially in the pharmaceutical and agricultural chemical industries, developed internal registries very early. One result of these registries was that new compounds with potentially marketable activity were always referred to by the registry number – typically a two letter company code followed by 4-5 numerals -- at least until they acquired a trade name. One side effect that began appearing in the 1960s was the appearance of two sets of numbers: one, a serial number, used only internally, and the second, a randomly assigned number for use on the outside. The reason was to not allow the competition to infer the level of the research activity.

Development of registry numbers for chemicals:

Abstracting and indexing organisations that made information on scientific publications available began to see the advantages of registry numbers for not only codifying information in their files, but also to allow another means of searching the files. Prominent examples are the Chemical Abstracts Service (CAS) Registry System, the Derwent List of Registry Compounds, and the Beilstein file. Every compound in the Beilstein file has a Beilstein Registry Number (BRN), as well as Beilstein System. The BRN is strictly a serial number, but the Derwent Registry Numbers (DRN) implicitly contain information on chemical composition and structure. However, they are assigned only to about 2000 compounds commonly encountered in the patent literature.

The Structure of the Chemical Abstracts Service – CAS numbers:

As defined by Chemical Abstracts Service – CAS, a division of the American Chemical Society, a registry system is an inventory of chemical substances with means of inputting, processing, searching, retrieving, and outputting information about the substances. Although the fundamental representation of substances is by structure, structure representation can be accomplished by:

- 1) nomenclature,
- 2) linear notation (e.g., Wiswasser line Notation WLN, SMILES code),
- 3) fragmentation codes (e.g., GREMAS, Derwent), and
- 4) connection tables (e.g., Morgan, DARC).

CAS chose the Morgan connection table method for the CAS Registry File.

The CAS Registry System was originally designed as a labour saving device in support of the indexing effort used to prepare the Chemical Abstracts database of scientific chemical literature. Prior to the advent of CAS Registry in 1965, with the exception of a list of ca. 2500 common chemicals, there was no good way to

determine if a compound had already appeared in the database. Each potentially new compound had to be drawn and named and the name compared to the list of index names. Since the advent of Registry II in 1968, compounds to be indexed undergo name and structure matching procedures against the Registry File. Those found are henceforth referred to by CAS Registry Number (CASRN). Those not found by either method are added to the file as new compounds and a Registry Number is generated.

The design, implementation, and performance of the CAS Registry System have been well documented over the years.

The format of CAS numbers:

The format of CAS Registry Numbers is intentionally unique. The number is a serial number; the next available number is assigned to the next new compound to be entered into the file. For example, formaldehyde – CASRN 50-00-0, or compound number "5000" -- is serially the second compound in the file. The last number is an algorithmically assigned check character. The format is from two to six numerals followed by a hyphen, followed by two numerals, followed by a hyphen, and concluding with one numeral. As indicated, 50-00-0, formaldehyde, is compound 5000, and 1746-01-6, dioxin, is compound 174,601. Note that the format makes the numbers more readily identifiable even at a glance.

1- butene is provided as an example of a CAS Registry File record:

L1 ANSWER 1 OF 1 REGISTRY COPYRIGHT 1999 ACS

RN 106-98-9 REGISTRY

CN 1-Butene (8CI, 9CI) (CA INDEX NAME)

OTHER NAMES:

CN .alpha.-Butene

CN .alpha.-Butylene

CN 1-Butylene

CN Butene-1

CN Ethylethylene

FS 3D CONCORD

DR 1735-75-7, 54366-07-3, 33004-02-3

MF C4 H8

CI COM

LC STN Files: AGRICOLA, ANABSTR, APILIT, APILIT2, APIPAT, APIPAT2, BEILSTEIN*, BIOBUSINESS, BIOSIS, CA, CAOLD, CAPLUS, CASREACT, CEN, CHEMCATS, CHEMINFORMRX, CHEMLIST, CBNB, CHEMSAFE, CIN, CSCHEM, CSNB, DETHERM*, DIPPR*, EMBASE, GMELIN*, HODOC*, HSDB*, IFICDB, IFIPAT, IFIUDB, MEDLINE, MRCK*, MSDS-OHS, NIOSHTIC, PIRA, PROMT, SPECINFO, TOXLIT, TOXLIT, TRCTHERMO*, TULSA, USPATFULL, VTB

(*File contains numerically searchable property data)

Other Sources: DSL**, EINECS**, TSCA**

(**Enter CHEMLIST File for up-to-date regulatory information)

7746 REFERENCES IN FILE CA (1967 TO DATE)

114 REFERENCES TO NON-SPECIFIC DERIVATIVES IN FILE CA

7758 REFERENCES IN FILE CAPLUS (1967 TO DATE)

11 REFERENCES IN FILE CAOLD (PRIOR TO 1967)

The Registry Number is shown, followed by nomenclature – first index name, then synonyms, deleted Registry Numbers (if any), molecular formula, source of registration, files containing this Registry Number (locator field), structure, the existence of abstract references in the CAOLD file (if any), and the current number of references in the CA and CAPLUS files.

There are a number of reasons for deleted Registry Numbers. One is a different source of registration that lead to duplication of CASRN for the same substance. Another was an early attempt to provide additional Registry Numbers, which would link compounds similar in stereochemistry. If these were encountered earlier, especially in CAS files, they are preserved in the database and a search for them will retrieve the current compound record in the Registry File.

A list of chemical names and Registry Numbers is the Registry Handbook—Common Names, which is published on microform. The complete list of Registry Numbers also appears in print as the Registry Handbook.

The CAS Registry Numbers:

Even in the current Registry System, a number of CASRN exist that have no references in bibliographic or data files. They merely register substances (usually less defined compositions or mixtures) that appear on one or more of the national regulatory lists, like the Toxic Substances Control Act of the USA (TSCA) Inventory or the European Inventory of Existing Chemical Substances (EINECS) of the European Union. In the process of preparing such inventories of commercially available substances the Chemical Abstract Service was involved and supplemented CASRN number for those substances that had no CASRN because they had not appeared in the chemical literature in the past.

Not only has the CAS Registry System become the cornerstone of the CAS indexing process, but also CAS Registry Numbers are being used for a number of other purposes. They appear in almost 60 files on the STN system, several of which are produced by CAS.

CASRN are used in sales (chemical catalogues, including Aldrich, Merck), transportation (including export/import), regulatory agency reporting, and disposal. In fact, they are required in many of these cases, especially the latter.

At the beginning of the Registry Systems era, CAS Registry numbers appeared in the abstract as well as in the index phrases. However, that practice has been discontinued. Since 1987, author inspired names have been provided for Registry Numbers in CA File index term phrases.

Assignment of CASRN by the CAS Registry System is quite accurate. Less accurate assignments are usually the result of author supplied information that is insufficiently accurate. This is often the case with patents, which are often written by non-chemists.

Appendix II

Information for New Visitors (from WWW.CAS.ORG/faq.html)

This is a quick orientation and overview of CAS for new visitors to our web site. It is presented in the style of a **FAQ (Frequently Asked Questions)** sheet.

- What is CAS?
- What is a CAS Registry Number and What do Those Numbers Mean?
- Tell me briefly about CAS Products
- Does CAS offer any Free Services?
- Which CAS Product is Right for Me?
- How Much Will It Cost?
- How Can a Publication Get Abstracted/Indexed by CAS?
- Is There a Listing of the Journals Covered by CAS?
- Does CAS have a Privacy Policy?
- Who Do I Contact if I Have a Question or Problem?

QUESTIONS: What is CAS?

ANSWER: CAS (Chemical Abstracts Service) is a division of the American Chemical Society. We are located in Columbus, Ohio. CAS is the producer of the largest and most comprehensive databases of chemical information. Our principal databases are Chemical Abstracts (CA) which contains over 18 million document records from the chemical journal and patent literature, and Registry which contains over 33 million substance records.

CAS also operates the STN International online service with partner organizations in Europe and Asia. STN provides access to nearly 200 databases covering topics such as science, technology, patents, and business information.

QUESTIONS: What is a CAS Registry Number?

ANSWER: CAS Registry Numbers (often referred to as CAS RNs or CAS Numbers) are unique identifiers for chemical substances. A Registry Number itself has no inherent chemical significance but provides an unambiguous way to identify a chemical substance or molecular structure when there are many possible systematic, generic, proprietary, or trivial names.

CAS Registry Numbers are used in many other public and private databases as well as chemical inventory listings and of course are included in all CAS produced databases.

QUESTIONS: Tell Me Briefly about CAS Products

ANSWER: CAS offers print, CD-ROM, online, desktop, and web based access to databases covering science, engineering, technology, patents, business information, and much more. These products are designed to accommodate a wide range of information needs, whether you are an infrequent searcher who only needs a few quick answers, or a professional searcher who requires a more powerful and comprehensive set of search tools.

CAS and STN products are fee based and generally require one to set up an account.

Read on to see which CAS product might be right for you.

QUESTIONS: Does CAS offer any free services?

ANSWER: There are free demo versions of CAS products including STN Easy, CA on CD, SciFinder, and SciFinder Scholar.

In the STN online system there are "Learning file" versions of CAS' principal databases CA and Registry. Learning files are subsets of the complete databases and users are only charged a small connect hour fee. There are no search or display charges when using Learning files. Contact CAS Customer Service for details.

In addition, CAS also provides a free news service from Reuters, which reports the latest business news on the chemical, pharmaceutical, biotechnology, and related industries.

QUESTIONS: Which CAS Product is Right for Me?

ANSWER: This is a basic yet oftentimes difficult question to quickly answer for new users. It depends on many factors such as how comprehensive your search and answers need to be, your familiarity and experience with online searching, the medium you wish to search in, how frequently you need to search, and of course cost is also a consideration.

If you are still unsure about things after reading this brief overview, a CAS Customer Service representative will be happy to help you. Or, you can try the CAS Product Finder. Simply answer few questions and we will point you towards the product that best matches your answers!

- SciFinder is a desktop tool suitable for both professional searchers and research scientists. The point and click interface means you don't have to be an expert searcher to use SciFinder. Many organizations around the world use SciFinder to give their scientists direct access to CAS databases. SciFinder packages and subscriptions can be tailored to smaller groups as well.
- SciFinder Scholar is similar to SciFinder except it is for use in academic settings only.
- STN International is the access point to a comprehensive collection of databases covering science, technology, patents, engineering, and business information. It is a command based online service and at least some training in the STN search language is recommended in order to benefit from the precision and power of the system. The STN Express with Discover software package will help make STN searching more efficient and productive.
- STN on the Web brings all the power and functionality of the dial-up STN to a web based system. Structure searching capabilities are included.
- STN Easy is a web based interface to over 50 of the most popular databases available on STN. There is no command language to learn and so this service is very much suited to both infrequent searchers and those who need a simpler and more guided information resource.
- CAS offers CD-ROM and Print products suitable for both institution (comprehensive collections) and individual (current awareness) use. Those involved with regulatory compliance might be especially interested in the 1 disc National Chemical Inventories collection.

QUESTIONS: How Much Will It Cost?

ANSWER: Costs depend on the product you are using. Products like STN, STN Easy and Chemical Patents Plus are "pay as you go" services with no subscription or monthly fees. The full STN Price List is available as a PDF file. Follow the links for details on STN Easy costs and Chemical Patents Plus prices respectively. SciFinder pricing is based on the number of users and the extent of access required. You should contact CAS Customer Service for further information. Prices for other products are listed in the CAS Catalog.

QUESTIONS: How Can a Publication Get Abstracted/Indexed by CAS?

ANSWER: Publishers interested in having a journal considered for inclusion in Chemical Abstracts should send either a copy of the first issue (if it is a new publication) or 3 different and recent issues if it has already been published for some time to:

Chemical Abstracts Service - CAS
Acquisitions / Evaluation Dept.
P.O. Box 3012
Columbus, OH 43210
U.S.A.

Electronic journal publishers can send URL pointers to:
Email:ctsacq@cas.org

We will be delighted to have our selector team of chemists determine whether the journal meets our editorial criteria for abstracting and indexing in CA.

QUESTIONS: Is There a Listing of Journals Covered by CAS?

ANSWER: There is a list of the key/core journals covered in CAplus. This listing includes both the complete and abbreviated journal titles, the CODEN, and frequency of publication. Note however, that this is not the complete list of journals covered by CAS.

QUESTIONS: Does CAS Have a Privacy Policy?

ANSWER: Yes. The full statement is available at <http://www.cas.org/privacy.html>

QUESTIONS: Who Do I Contact when I have a Question or Problem?

ANSWER: You have several options depending on the nature of your question or concern. Use the following as a guide;

- Contact CAS Customer Service for general questions about CAS or STN products. This would include account set up and status, costs, ordering information, obtaining user documentation, etc. CAS Customer Service is open from 8am-8pm Eastern Time (Monday-Friday)
 - Contact the CAS/STN Help Desk when you have technical questions about using CAS products. Help Desk hours are also 8am-8pm Eastern Time (Monday-Friday).
 - Contact the CAS Document Detective Service (CAS DDS) when you need to order a copy of a journal article or patent
 - Contact CAS Client Services when you need us to run a search for you, or if you need confirmation/assignment of CAS Registry Numbers and chemical names
 - For enquiries outside North America be sure to check the list of official CAS/STN representatives who can answer CAS and STN related questions.
 - We would also highly recommend trying the Site Search feature of this web site. It doesn't search the fee based databases for references and Registry Numbers, etc., but it is a useful way to find content on the web site. The Site Search function supports both concept and keyword searching (with Boolean operators).
-