the Cambridge Centre, or through accredited data centers on a regional basis.

Accredited data centers are being established in various parts of the world and are intended to serve the needs either of a large research institution or of a scientific community on a geographic basis. An accredited center may lease the whole or a part of the data base and use it to provide a variety of services subject to certain lease restrictions.

ACKNOWLEDGMENT

We thank the Office for Scientific and Technical Information, Department of Education and Science for their support of the activities of the Crystallographic Data Centre, and individual staff members for their advice and helpful discussion. Space for the Centre was provided by the University of Cambridge and computing facilities by the Institute of Theoretical Astronomy (IBM 360/44) and the Cambridge University Computing Service (IBM 370/165). T. Scott of the Medical Research Council Computing Service took a large part in the development of the evaluation program SYSMOL. We are indebted to the following for their assistance during preliminary abstracting, data input, and checkout processes: Stella Weeds, Ann O'Brien, Eileen Doyle, Anne Town, Bernice Roberts, Adele Boskey, Maryla Chiechanowicz and D. Sales.

LITERATURE CITED

(1) Kennard, O., Watson, D. G., and Town, W. G., *J. Chem. Doc.* **12,** 14 (1972).

(2) Kennard, O., Watson, D. G., "Molecular Structures and Dimensions," Vols. 1, 2, 3, and 4 (with Town, W. G.) published for the Crystallographic Data Centre, Cambridge, and the International Union of Crystallography by N. V. A. Oosthoek's Uitgevers Mij, Utrecht, 1970, 1972, 1973

(3) Allen, F. H., Kennard, O., Motherwell, W. D. S., Town, W. G., Watson, D. G., Scott, T., and Larson, A. C., to be published.

(4) Kennard, O., Watson, D. G., Allen, F. H., Isaacs, N. W., Motherwell, W. D. S., Pettersen, R. C., and Town W. G., "Molecular Structures and Dimensions, Vol. A1, Interatomic Distances 1960–65," published for the Crystallographic Data Centre, Cambridge and the International Union of Crystallography by N. V. A. Oosthoek's Uitgevers Mij, Utrecht, 1973.

(5) "Numerical Data and Functional Relationships in Science and Technology," III, Vol. 5a, 5b, Eds. K-H and A. M. Hellwege, Berlin Springer, 1971.

A Materials Index—Its Storage, Retrieval, and Display*

CAROL Z. ROSEN

Information Services Division, American Institute of Physics, 335 East 45th St., New York, N. Y. 10017

Received April 19, 1973

An experimental procedure for indexing physical materials based on simple syntactical rules was tested by encoding the materials in the journal, *Applied Physics Letters,* to produce a materials index. The syntax and numerous examples together with an indication of the method by which retrieval can be effected are presented.

In today's research and development, there are myriad physical materials available to the scientist or engineer for his work. A materials index to the scientific journals can be of great assistance in furthering the scientist's knowledge and accelerating his achievements. Our goals have been to derive a simple scheme compact enough to minimize the number of material entries per article and general enough to encompass the full range of materials and devices discussed in the present and future literature. We have tried to make the coding of the materials easy to learn, simple to recognize, and based, wherever possible, on codes which already exist in the literature. The feasibility of computer retrieval of items on the index played a large role in the selection of these codes.

It is anticipated that the Materials Index can be produced by computer-based photocomposition directly from a computer file containing a list of say, *Applied Physics Letters* (APPLA) articles and their materials. This file in turn could be merged with a file holding the Physics and Astronomy Classification Scheme (PACS)[1] information for these APPLA articles. PACS is currently used in organizing the material appearing in *Current Physics Titles* (CPT), as well as the three monthly journals *Current Physics Advance Abstracts* (CPAA) published by the American Institute of Physics.

To illustrate this point, suppose a user is interested in Kondo-materials, in particular, Pd-Ni and Pd-Mn. The materials index encoding scheme (MIES)-PACS combination could link these two specific alloys with papers about order-disorder phenomena, localized electronic states, resistivity, electronic specific heat, and magnetic impurity interactions. From this material—subject merger, the user will get a picture of the current state of research on these two Kondo alloys. In general, this system can provide detailed retrieval and display for all subjects of physics.

MATERIALS INDEX ENCODING SCHEME (MIES)

Tables I and II display the full glossary for the encoding scheme. Table I lists the "Type" symbols, along with their physical meaning, and a material example. Table II describes each "Delimiter" along with descriptive examples. The "Type" symbol describes the material's physical state or the physical process it underwent such as implan-

*Work supported by NSF Grant GN864

Table I. Glossary* of Types

(Type symbols alphabetized)

No.	Type	Character	Name (alphabetized)	Physical Meaning	Material Example
1	↓	binary	adon	sorption materials	Si ↓ He or Si adon He or He sorbedon Si
2	—	binary	alloy	alloyed materials	Al-Cu or Al alloy Cu
3	@ g	unitary	at gas	at gas state	I_2 @ g or I_2 at gas
4	@ l	unitary	at liquid	at liquid state	^4He @ l or ^4He at liquid
5	@ p	unitary	at plasma	at plasma state	Ar @ p or Ar at plasma
6	@ s	unitary	at solid	at solid state	H_2 @ s or H_2 at solid
7	•	binary	bind	molecular binding	$CuSO_4 \cdot 5H_2O$ or $CuSO_4$ bind $5H_2O$
8	:	binary	dopant	material containing impurities	CdS : Ag or CdS dopant Ag or Ag doped CdS
9	|	binary	film	material containing films	Cu | Au or Cu with Au film or Au on Cu
10	interface	binary	interfacing surfaces	material consisting of interfacing (or juxtaposed) substances	GaAs interface $LiNbO_3$
11	†	unitary	laser	lasing material	CO_2 † or CO_2 laser
12	<	binary	mimpact	mass impact materials undergoing irradiation e.g., bombardment, channelling implantation and sputtering	Eu < Ar^+ or Eu mimpact Ar^+ or Ar^+ irradiate Eu or Eu irradiated by Ar^+
13	>	binary	pimpact	photon impact materials undergoing radiation	Cu > [CO_2†] or Cu radiated by CO_2 laser or CO_2 laser radiating Cu
14	→	binary	react	Chemical reaction	$(CS_2, O_2) \to (CO_2, SO_2)$ or $2CS_2 + 5O_2 \to 2CO - 4SO_2$
				Nuclear particle reaction	$\gamma p \to \pi^0 p$
15	←	binary	send	material resulting from diffusion or migration process	Au ← Cr or Au send Cr or Cr send to Au

*The general form of an entry is *Primary Material Type Secondary Material.*

Table II. Glossary of Delimiters

No.	Delimiter	Name	Physical Meaning	Material Example
1	,	comma	presence of concommitant materials within list	CdS: (Ag, Cu) *or* CdS dopants Ag *and* Cu
2	;	semicolon	presence of exclusive alternates within list	CdS: (Ag; Cu) *or* CdS dopant Ag *and* CdS dopant Cu
3	( )	parentheses	indicates list of materials	CdS: Ag parentheses not necessary CdS: (Ag, Cu) parentheses needed CdS: (Ag; Cu) parentheses needed
4	[ ]	square brackets	indicates imbedded type	[Si:P] — ^{40}Ar also Cu > [CO_2†]

tation, diffusion, irradiation, lasing, etc. The "Character" column serves to differentiate the "Types" according to their unitary or binary nature.

This scheme thus represents a method for encoding materials onto an index. Fundamental to the scheme is the basic syntax of primary term list, type symbol, and secondary term list along with key delimiters. The Backus normal form[2] of this syntax is given in Table III. This table provides a concise summary of the entire scheme as regards the format of the items. Several points are noted here to clarify Table III. The definition of a ⟨Basic-Material⟩ is satisfactory for the recognition of a material, but could not be used to generate formulas of materials. Numbers and letters may also occur in subscripted and superscripted form; b means blank space.

Having cast this encoding scheme into Backus normal form, it is possible to take advantage of the body of work on syntax-directed recognition[3] for programming purposes. This scheme has proved sufficiently flexible to allow us to encode the materials studied in the articles published in APPLA from January 1971 up through the present time.

MATERIAL STORAGE AND RETRIEVAL BY COMPUTER

The materials index for a single journal over an extended time period can be generated from the input shown in Table IV. The materials are alphabetically listed for each article. The article number—e.g., 0001—is used to link every material to its article.

This input tape is sorted by computer into the desired alphabetized material list illustrated in Table V.

To devise a retrieval scheme we examine the formulation of the codes as defined in Tables I–III. In general, the codes consist of primary material or a primary list of materials followed by a type followed by a secondary material or a secondary list of materials. For example, consider the problem of retrieving all papers dealing with the material formed from ^3He adsorbed on TiO_2. Suppose the material is listed in an article under the code TiO_2 ↓ (Ar; ^3He; ^4He; N_2). We have

Primary Material = TiO_2
Type = ↓
Secondary List = (Ar; ^3He; ^4He; N_2)

Table III. Backus Normal Form of Syntax for Materials Index Encoding Scheme

```
<MATERIAL> ::=<LIST>|<LIST> <TYPE>|<TYPE> <LIST> |<LIST> <TYPE> <LIST>
<LIST> ::=<SIMPLE-MATERIAL>|(<SIMPLE-MATERIAL> <SEPARATOR> <SIMPLE-MATERIAL-LIST>)
<SIMPLE-MATERIAL-LIST> ::=<SIMPLE-MATERIAL>|<SIMPLE-MATERIAL> <SEPARATOR> <SIMPLE-
                                                                 MATERIAL-LIST>
<SIMPLE-MATERIAL> ::=<BASIC-MATERIAL> [<MATERIAL> ]
<BASIC-MATERIAL> ::=<BASIC-MATERIAL> |<BASIC-MATERIAL> <FORMULA-SIGN>
<TYPE> ::=<AT-TYPE>|—|↓|:|•|←|→|<|>|•|interface|†
<AT-TYPE>::=@<LETTER>
<SEPARATOR> ::= ;|,
<LETTER> ::=A|B .....|S|T|U|V|W|X|Y|Z|a|b|c|.... z|A|B|Γ|Δ|••••|Ω|α|β|γ|δ|•••• ω
<NUMBER> ::=1|2|3|4|..... 9|0
<FORMULA-SIGN>::=<LETTER>|<NUMBER>|-|+|-|+|'|(|)|b|.|'|'|*|_
```

Table IV. Input Data File

Material A	0001
Material B	0001
0001 Title, Authors, Volume, Page (Year)	
Material A	0002
Material B	0002
Material C	0002
0002 Title, Author, Volume, Page (Year)	
etc.	

Table V. Alphabetized Material List

Material A1	0006, 0007, 0009, ... , 8001
Material A2	0003, 0025, 0032, ... , 9999
Material B1	0041, 0060, 0073, ... , 7987
Material B2	0004, 0032, 0095, ... , 600
Material B3	0025, 0108, 0233, ... , 7900
·	
·	
·	
·	
Material Zn	0009, 0094, 0853

Retrieval of this paper could be accomplished by the following algorithm which we sketch below.

1. *Search* Primary Material. If TiO_2 absent stop, if TiO_2 present go to Type.

2. *Search* Type. If Type ↓ absent stop, if Type ↓ present go to secondary list.

3. *Search* Secondary List. If ^3He absent stop and continue to next article, or print negative result.

If ^3He is present then stop and print the article or hold it in storage and print all articles with material TiO_2 ↓ ^3He together after the search through the entire store is completed.

Table VI. First Page of the Materials Index to *Applied Physics Letters,* Volume 21 (1972), as Prepared for Publication

Papers are arranged in ascending page number order. The symbols after the title of the paper indicate the general emphasis of the paper: (T)—Theoretical, (E)—Experimental (T/E)—Theoretical and Experimental. They are listed under the materials used.

The commas in the materials listed below denote the concommitant presence of materials within a list; e.g., CdS :(Ag, Cu) indicates CdS *doped* with Ag and Cu. The semicolons denote the presence of exclusive alternates within a list; e.g., CdS :(__; Ag; Cu) indicates three distinct materials, viz., *undoped* CdS and CdS *doped* with Ag and CdS *doped* with Cu. Parentheses indicate a list of materials as shown above. Square brackets indicate a compound structure; e.g., [YAG : Nd] laser indicates a Nd-*doped* YAG *laser*. The type symbol | indicates *film* and the material to the left of type | is the substrate and the material to the right of type | is the film. The material NaCl | Au indicates an Au *film* on a NaCl substrate. The type symbol ← denotes *diffusion and migration*. The material Pyrex ← He indicates He *diffused* into Pyrex. The type synbol < is used to describe an *irradiation* process, e.g., *bombardment*, *channeling* and *implantation*. The material [n-Si : P] < B^+ indicates B^+ *implanted* into target [n-Si : P]. The type symbol > is used to describe a *radiation* process. The material [| chalcogenide glass] > [Ar laser] denotes an [Ar laser] *radiating* onto a [chalcogenide glass film] target. For both the *irradiation* < and the *radiation* > processes the material to the left of the type symbol is the target.

• APAPA (*p*-methoxybenzylidene-*p*′-aminophenylacetate) NLC

Dynamic scattering life in the nematic compound *p*-methoxybenzylidene-*p*′-amino phenyl acetate as influenced by current density (E)—Alan Sussman. 126

• (Ag;Al;Bi;Cr;Cu;Ni;Sn) on SiO_2

Interaction of ultrasonic surface waves with conduction electrons in thin metal films (E/T)—P. Bierbaum. 595

• [(Ag;Al)–Cu] > [[glass : Nd] laser]

Rapid cooling by laser melt quenching (E)—W. A. Elliott, F. P. Gagliano, G. Krauss. 23

• [AgBr on NaCl]

Guided-wave propagation at 10.6 μm in silver bromide thin films (E)—J.H. McFee, J.D. McGee, T.Y. Chang, V.T. Nguyen. 534

• (AgCl, H_2O) > [(He, Ne) laser]

Mass motion as observed by light-beating spectroscopy (E)—N. Ben-Yosef, S. Zweigenbaum, A. Weitz. 436

• | Agfa-Gevaert 10E56

Underwater optical holographic interferometry (E)—C.D. Johnson, G.M. Mayer. 369

• | Agfa Gevaert 10E75

Time-resolved differential holographic microscopy (E)—H.M. Presby. 31

• (air;Ar;CH_4;CO;CO_2;He;Ne;O_2) > [CO_2 TEA laser]

cw gas breakdown in argon using 10.6 μm laser radiation (E)—Douglas L. Franzen. 62

• [air <u>interface plasma</u> dielectric] > laser

Morphological asymmetry in laser damage of transparent dielectric surfaces (E)—N.L. Boling, G. Dubé, M.D. Crisp. 487

• [air < e^-] laser

Computed secondary-electron and electric field distributions in an electron-beam-controlled gas-discharge laser (T)—R.C. Smith. 352

• Al

X-ray diffraction study of single crystals undergoing shock-wave compression (E)—Quintin Johnson, Arthur C. Mitchell, L. Evans. 29

Stress-pulse propagation in solids: a closer look at dispersion (E)—M.P. Felix, A.T. Ellis. 532

Visual display of fatigue damage by means of exoelectron emission (E)—William J. Baxter. 590

• Al <u>interface Teflon FEP</u> Al

Condenser-transducer array for acoustical holography (E)—A.K. Nigam, G.M. Sessler. 229

• Al > [[CO_2 < e^-] laser]

Momentum transfer and plasma formation above a surface with a high-power CO_2 laser (E)—A.N. Pirri, R. Schlier, D. Northam. 79

• [Al | (Duco cement;RTV602;silicone)] > [[glass : Nd] laser]

Laser-induced stress-wave and impulse augmentation (E)—J.D. O'Keefe, C.H. Skeen. 464

• [(Al, Corning 7059 glass) on Vycor] > [(He, Ne) laser]

Fundamental transverse electric field (TE_0) mode selection for thin-film asymmetric light guides (E/T)—Yasuharu Suematsu, Mitsuari Hakuta, Kazuhito Furuya, Kazuhiro Chiba, Ritsuo Hasumi. 291

• [| ($Al_{0.2}Ga_{0.8}As$, GaAs, $Al_yGa_{1-y}As$)] laser

Differential *I/V* of heterostructure correlates with laser threshold (E)—D.L. Rode, L.R. Dawson. 90

• Al_2O_3

High-speed photography of surface flashover of solid insulators under impulse voltages in vacuum (E)—J.D. Cross, K.D. Srivastava. 549

• Al_2O_3 | Al

Internal photoemission in sapphire substrates (E)—C.R. Viswanathan, R.Y. Loo. 370

• Al_2Pt

Platinum silicide-aluminum Schottky diode characteristics (E)—H.H. Hosack. 256

• Ar laser

Underwater optical holographic interferometry (E)—C.D. Johnson, G.M. Mayer. 369

• Ar > [CO_2 TEA laser]

cw gas breakdown in argon using 10.6-μm laser radiation (E)—Douglas L. Franzen. 62

• (Ar, <u>CO</u>, N_2, O_2) laser

Electrical CO-mixing gas-dynamic laser (E)—H. Brunet, M. Mabru. 432

• (Ar, <u>Cu</u>, He) laser

Supersonic electrical-discharge copper vapor laser (E/T)—G.R. Russell, N.M. Nerheim, T.J. Pivirotto. 565

• (Ar;He) plasma

Large quiescent collisionless high-frequency discharge plasma (E)—Earl R. Ault, K.R. MacKenzie. 9

In general processing by computer is much the same for all materials encoded in this Backus normal form.

MATERIALS INDEX DISPLAY TO APPLA (1971)

In compiling the materials index, we observe the following rules for alphabetizing. Alphabetize first on primary material then on type then on secondary material.

The rules for alphabetizing on the primary material are:

1. Elements are ordered alphabetically—e.g., Al, Ar, As.
2. Negative ions precede positive ions, e.g., H^-, H^+.
3. Multiply ionized materials are arranged in order of increasing valence number—e.g., Fe^{+2}, Fe^{+3}.
4. Materials such as alcohol are listed alphabetically—e.g., Al, alcohol, Ar, As.
5. Materials such as α-quartz and α-particles are alphabetized according to the leading word alpha—e.g. Al, α-particles, α-quartz, As, Be, β-particles, Bi.
6. Compounds are listed according to first element then ordered according to number of first elements—e.g., CH_4, C_2H_8, C_6H_6.
7. p- and n-type semiconducting materials such as n-Ge and p-Ge are alphabetized under the name of the undoped substance. As for example n-Ge would be listed under g. Symbols p, o, m, denoting para, ortho, and meta positions on aromatic ring compounds or ortho and para hydrogen, etc., are ordered in a similar manner. Wherever possible, we follow the rules of alphabetization set forth in *The Condensed Chemical Dictionary*.[4]

   All n, p, m, o symbols and numbers have a hyphen separating them from their compounds as shown in the following 2 examples

   a. n-GaAs

   b. 4, 4′-di-*n*-heptyloxyazobenzene:vanadyl acetylacetonate

   Example b is alphabetized under d in accordance with the dictionary's rules.
8. The ordering on a primary list is alphabetical. In the case of lasers, the lasing material is underlined if it is not the leading term in the list.

The ordering rules imposed on the secondary materials are the same as the rules for the primary materials with the exception that the ordering of multiple thin films listed within the secondary string is significant. In this case, the films are not sorted alphabetically but in order of their relative position to the substrate. Consider the example of the secondary list for the material formed by a V film upon an Er film upon a Sc film upon a Kovar substrate. This material is coded as Kovar | (Sc, Er, V).

The "Type" codes and their alphabetized names are listed in Table I along with columns for a brief physical meaning and material example. The purpose of these names is to allow for the alphabetic ordering of the "Type" codes and to provide a means by which the materials index can be displayed without symbols.

Substances comprised solely of primary coded materials are alphabetized first. Primary strings containing commas only precede those strings containing semicolons. Strings with both delimiters are ordered after these. The primary substances are followed by those materials which are of the form primary material and type. They in turn are followed by materials of the form primary material, type, secondary material. These in turn are followed by materials in which the primary or secondary or both lists contain imbedded type symbols.

Table VI is a display of a portion of the *Materials Index* to APPLA, volume 21, numbers 1-12 (1972). The materials are alphabetically listed in ascending page order as they would appear in the actual index. Some of the type symbols such as †, @ p are written out for ease in reading. Note again that each type symbol within a primary or secondary list is enclosed within square brackets. This delimiter serves as a signal to the computer that there is an enclosed type symbol within the string and will facilitate the ensuing search. Brackets and type symbols can be further suppressed as desired.

CONCLUSIONS

We have presented an encoding scheme for materials (MIES) and illustrated its application to materials discussed in articles appearing in *Applied Physics Letters*. This experiment tests the scheme because of the wide variety of materials which appear in this journal. The scheme's ability to be merged with a subject index, such as PACS or *Nuclear Science Abstracts*,[5] makes MIES useful for in-depth retrieval.

LITERATURE CITED

(1) PACS-Physics and Astronomy Classification Scheme published by the American Institute of Physics, AIP-ID-71B.

(2) Revised Report on the Algorithmic Language ALGOL 60, P. Naur, Ed., reprinted in "Programming Systems and Languages," S. Rosen, p. 79, McGraw-Hill, New York, N. Y., 1967.

(3) See for example T. E. Cheatham and K. Sattley, Syntax Directed Compiling, S. Rosen, *op. cit.* p. 264.

(4) *The Condensed Chemical Dictionary* (Seventh Edition) edited by Arthur and Elizabeth Rose, Reinhold Publishing Corp., New York, N. Y., 1966.

(5) Nuclear Science Abstracts, published by USAEC Technical Information Center, Oak Ridge, Tenn.