Bringing Aristotle into the 20th Century:

Definition-Oriented Concept Analysis in a Terminological Knowledge Base

by

Karen Eck

School of Translation and Interpretation
University of Ottawa

Under the supervision of
Ingrid Meyer, Ph.D.
School of Translation and Interpretation
and
Douglas Skuce, Ph.D.
Department of Computer Science

Thesis submitted to
the School of Graduate Studies and Research
of the University of Ottawa
in partial fulfilment of the requirements
for the degree of M.A. (Translation)

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-315-95948-7

Canada

That it is more difficult to establish than to overthrow a definition, is obvious....

—Aristotle (384–322 B.C.)
Topica VII

Acknowledgements

There were times when I didn't think I'd finish this magnum opus (all things being relative), but I was lucky enough to have many kind and understanding people on my side, all of whom deserve my thanks:

· Dr. Ingrid Meyer, for her neverending guidance, patience, and encouragement. She has been more than a supervisor to me: she has been a mentor, and I don't think I'd be where I am today (with a finished thesis and a job) without her.

· Dr. Douglas Skuce, for allowing me to use the CODE system and for his input and encouragement.

· The people behind the software, especially Ken Iisaka and Timothy Lethbridge. I'm sure there are very few Master's students in the world with their own personal programmers, and I was very lucky to have them at my side.

· François Allard and Michel Racine, for all their patience and help in getting my CODE graphics into WP5.1.

· Dr. Geneviève Mareschal, for making the initial suggestion which led to the creation of the Characteristics Comparison Matrix, and both Dr. Mareschal and Dr. Roda Roberts for commenting on very early drafts of the thesis.

· Dr. Emory Fortin, Dr. Claude Rolland, and especially Dr. Doug Beckett, for taking time out of their busy schedules to act as my laser subject-field experts.

· David R. Miller, for being kind enough to read through my final draft.

· The Provincial Government of Ontario and the University of Ottawa School of Graduate Studies and Research, for providing me with funding through the Ontario Graduate Scholarship and the University of Ottawa Supplementary Scholarship, respectively.

· BodyWorks, for keeping me in relative shape and sanity.

· My family (Mom, Dad, Kevin, Jamie, Nana, and Spike), for tactfully enquiring how things were going when they knew they really weren't and for always showing support, even when I'm pretty sure they had no idea what I was talking about.

· Al, Bruce, Ellen, Michele, and especially Dave, for not smirking whenever I'd declare "I'm working on it!" and for putting up with my bad moods when I actually was.

· Daniel, for his patience, sympathy, love, translation skills, and time management tips.

· Lynne (L.M.Y) Bowker, who may have been an ocean away, but who was always there for me. I couldn't ask for a better friend or (unlicensed) therapist.

Abstract

This thesis develops and tests a definition-oriented approach to concept analysis. In this approach, a computerized knowledge management system called CODE is used to help the terminologist carry out concept analysis specifically aimed at constructing intensional (Aristotelian) definitions for a terminological knowledge base (TKB).

The approach is developed by first exploring what the literature has to say about the intensional terminological definition (ITD) as a process, a product, and a tool. The primary conclusion that arises from this exploration is that ITDs for co-concepts must be consistent in content and form, and must clearly indicate the place of their respective concepts in the concept system. Second, this thesis looks at how a specific feature of the CODE system, the Characteristic Comparison Matrix (CCM), can be enhanced to facilitate definition-oriented concept analysis. The result is a powerful, flexible feature that aids the terminologist in recording consistent characteristics and selecting characteristics for definition construction. Third, this thesis proposes enhancements to an existing general approach to concept analysis. These enhancements are based on definition research, knowledge of CODE, and introspection, supported by previous experience performing general concept analysis with the COGNITERM Project. The result is an approach which allows the terminologist to carry out concept analysis in a methodical manner, using CODE to create good, consistent ITDs for a TKB.

This approach is tested by creating a TKB for the subject field of lasers (concentrating on 16 concepts), complete with ITDs. The approach works quite well, although it requires some minor modifications. The definitions are consistent and contain only that information necessary to distinguish co-concepts from one another and to indicate their place within a concept system. Although CODE is helpful in facilitating definition-oriented concept analysis, some enhancements to the technology would make it even more useful.

Résumé

La présente thèse a pour objet de proposer et de mettre à l'essai une approche à l'analyse notionnelle axée sur la définition. Cette approche fait appel au système informatisé de gestion des connaissances appelé CODE, qui aide le terminologue à effectuer une analyse notionnelle dans le but de créer des définitions caractéristiques, autrement appelées aristotéliciennes, qui pourraient s'intégrer à une base de connaissances terminologiques (BCT).

L'auteure examine en premier lieu l'état de la recherche afin de délimiter les contours de la définition caractéristique terminologique (DCT) à titre de processus, de produit, d'outil. Cette étude lui permet de conclure d'emblée que les DCT des notions coordonnées doivent être conséquentes au plan du contenu et du contenant et qu'elles doivent indiquer clairement la place de leurs concepts respectifs dans le système notionnel. En deuxième lieu, l'auteure se penche sur un attribut précis du système CODE, la Matrice des comparaisons de caractères (MCC), dans le but de déterminer dans quelle mesure cet outil facilite l'analyse notionnelle axée sur la définition. Enquête faite, on y voit un outil puissant, flexible, qui aide le terminologue à relever les caractères conséquents et à choisir ceux qui servent à la création des définitions. En troisième lieu, l'auteure propose des améliorations à l'approche existante à l'analyse notionnelle générale. Ces améliorations s'inspirent des recherches sur les définitions effectuées par l'auteure, sa maîtrise du système CODE et des connaissances qu'elle a acquises en procédant à des analyses notionnelles générales dans le cadre du projet COGNITERM. Il en résulte une approche qui permet au terminologue d'effectuer méthodiquement une analyse notionnelle en faisant appel au système CODE afin de mettre au point des DCT sûres et conséquentes qui sauront s'intégrer à une BCT.

L'auteure fait ensuite l'essai de cette approche en créant une BCT pour le domaine du laser (en se limitant à 16 notions) qui comprend toutes les DCT voulues. L'approche réussit fort bien, quoiqu'elle nécessite des modifications mineures. Les définitions sont conséquentes et ne contiennent que les informations qu'il faut pour distinguer les notions coordonnées les unes des autres et pour indiquer leur place à l'intérieur d'un système notionnel. Quoique le système CODE facilite grandement l'analyse notionnelle axée sur la définition, quelques améliorations de nature technologique le rendrait encore plus utile.

Table of Contents

List of Figures

List of Tables

Notational Conventions

The following notational conventions are used throughout this thesis:

1. Terms which are defined in the glossary appear in UPPER CASE the first time they are used in the body of the thesis:

   e.g. There are two types of SUPERCONCEPT: GENERIC CONCEPTS and COMPREHENSIVE CONCEPTS.

2. Terms denoting concepts are in **bold**:

   e.g. The concept **CD-ROM** inherits the characteristics of its generic concept, **compact disc**.

3. Terms denoting characteristic names are in ***bold and italics***.

   e.g. It is interesting to note that the characteristic name ***error correction*** is also a specialized concept.

4. Terms denoting characteristic values are *italicized*.

   e.g. Of course, it is also possible for a characteristic value, such as *layered error correction code*, to be a specialized concept.

   Italics are also used to show emphasis.

   e.g. ...while a LKB tries to capture what a native speaker knows about general lexical units, a TKB tries to capture what a native speaker *who is also a subject-field expert* knows about *specialized* lexical units (terms).

5. Terms denoting characteristics are a combination of ***characteristic name*** and *characteristic value* (see 3 and 4 above), joined by a colon.

   e.g. The concept **CD-ROM** has the characteristic ***error correction:****LECC (layered error correction code)*.

6. Terms denoting facets (used to describe characteristics) are a combination of facet name and facet value, joined by a colon. The facet name is in ***BOLD, ITALICS, AND UPPER CASE***, while the facet value is *ITALICS AND UPPER CASE*.

   e.g. The characteristic ***content:****audio* has the facet ***STATUS:****UC*.

Glossary

NOTES:

(1) The entries in this glossary consist of definitions intended to provide the reader with a general explanation of the important concepts discussed in this thesis.

(2) I would like to thank Lynne Bowker for providing several entries from her Master's thesis glossary [cf. Bowker 1992]. These entries, which are followed by the initials [LB], have been used (sometimes in slightly modified form) in order to provide consistent terminology for all documents derived from the COGNITERM Project.

(3) Any term appearing in UPPER CASE within a definition has its own entry in the glossary.

(4) Any term in this glossary will be in UPPER CASE in the body of the thesis the first time it appears.

AI: see ARTIFICIAL INTELLIGENCE (AI).

approach: a means of carrying out a given activity (e.g. CONCEPT ANALYSIS), characterized by a series of steps.

NOTE: I have used the term *approach* instead of other more scientific-sounding terms such as methodology, procedure, etc. because of the non-sequential nature of the "steps" to be followed, and because definition construction is truly part science, part art.

artificial intelligence (AI): the SUBFIELD of computer science concerned with designing intelligent computer systems (systems that exhibit the characteristics associated with intelligence in human behaviour, such as understanding language, learning, reasoning, solving problems, etc.). [LB]

attribute: a type of CHARACTERISTIC which is inherent to the CONCEPT being described, and which does not have a RELATION to other specialized CONCEPTS (e.g. the CONCEPT **fax machine** could have the ATTRIBUTES ***height:30 cm, width:40 cm, weight:20 kg***).

characteristic: a quality that distinguishes or identifies a CONCEPT. A characteristic has two components: a name and a value. For example, one of the characteristics of the CONCEPT **optical disc** is ***physical form:disc***. The characteristic name is ***physical form***, while the characteristic value is ***disc***. (ISO 1990a uses the term 'property' to refer to an OBJECT, and 'characteristic' to refer to a CONCEPT.) [LB] (See also: NECESSARY CHARACTERISTIC, N+S CHARACTERISTIC, TYPICAL CHARACTERISTIC.)

CODE (Conceptually Oriented Description Environment): a generic knowledge management system developed at the Artificial Intelligence Laboratory, Department of Computer Science, University of Ottawa.

COGNITERM: a TERMINOLOGICAL KNOWLEDGE BASE built using CODE within the framework of the COGNITERM Project. (The COGNITERM Project is a research project which uses CODE to explore how terminology and knowledge engineering can benefit each other. It is being carried out at the Artificial Intelligence Laboratory, Department of Computer Science, University of Ottawa.)

comprehensive concept: a SUPERCONCEPT in a PARTITIVE RELATION with several PARTITIVE CONCEPTS (a CONCEPT of which the whole is divided into parts). For example, the CONCEPT **bicycle** is a comprehensive concept relative to the PARTITIVE CONCEPTS **wheel, seat** and **frame.**

concept: a unit of thought constituted through abstraction on the basis of properties [see CHARACTERISTIC] common to a set of OBJECTS [ISO 1990a:1] (See also CO-CONCEPT, GENERIC CONCEPT, SPECIFIC CONCEPT).

concept analysis: the process of describing the CHARACTERISTICS of individual CONCEPTS and the RELATIONS that hold within CONCEPT SYSTEMS.

concept system: a set of CONCEPTS structured according to the RELATIONS between them.

co-concept: a CONCEPT in a hierarchical CONCEPT SYSTEM which ranks at the same level as one or more other CONCEPTS [ISO 1990a:2]. For example, **CD-ROM** and **CD-Audio** are CO-CONCEPTS since they are both GENERIC CONCEPTS of **compact disc; frame, wheel,** and **seat** are CO-CONCEPTS since they are all PARTITIVE CONCEPTS of **bicycle.**

definition: a statement about a CONCEPT which provides the essence of the concept, i.e. which distinguishes it from all other concepts.

definition-oriented approach to concept analysis: an APPROACH that allows a terminologist to carry out DEFINITION-ORIENTED CONCEPT ANALYSIS

definition-oriented concept analysis: CONCEPT ANALYSIS specifically for the purpose of constructing (or critiquing or modifying) definitions on the basis of CHARACTERISTICS.

differentia: TERM that denotes the set of N+S CHARACTERISTICS combined in an INTENSIONAL DEFINITION (see also GENUS).

NOTE: This term is also widely used with respect to definitions of general language expressions.

extension: the range of OBJECTS to which a CONCEPT applies.

facet:

generic concept: a SUPERCONCEPT in a GENERIC-SPECIFIC RELATION (i.e. a more general version of the SPECIFIC CONCEPT). For example, **optical disc** is a GENERIC CONCEPT of **compact disc.**

general approach to concept analysis: an APPROACH for carrying out GENERAL CONCEPT ANALYSIS

general concept analysis: CONCEPT ANALYSIS carried out without definition construction specifically in mind.

generic–specific relation: a hierarchical relation based on the partial identity of the intensions of GENERIC CONCEPTS, SPECIFIC CONCEPTS, and CO–CONCEPTS [ISO 1990a:3].

genus: the TERM that denotes the GENERIC CONCEPT presented in an INTENSIONAL DEFINITION.

hierarchical relation: a relation between two CONCEPTS which is established by the division of a SUPERCONCEPT into a SUBCONCEPT forming one or more levels, or by the reverse process [ISO 1990a:3]; types of HIERARCHICAL RELATION include the GENERIC RELATION and the PARTITIVE RELATION.

inheritance: in ARTIFICIAL INTELLIGENCE, a powerful technique that applies when CONCEPTS are arranged in a GENERIC–SPECIFIC HIERARCHY; it allows any CHARACTERISTIC of a given CONCEPT to be implicitly true for all SPECIFIC CONCEPTS of this CONCEPT, for all SPECIFIC CONCEPTS of those SPECIFIC CONCEPTS, etc. For example, if the CONCEPT **alcoholic beverage** has the CHARACTERISTIC ***alcohol content:up to 100%***, then its SPECIFIC CONCEPTS (**fermented alcoholic beverage** and **distilled alcoholic beverage**) will inherit this CHARACTERISTIC, as will in turn all their SPECIFIC CONCEPTS (e.g. **beer**, **wine**, **vodka**, **rum**, **whiskey**). [LB]

intension: the set of CHARACTERISTICS which constitutes a CONCEPT [ISO 1990a:2].

intensional definition: see INTENSIONAL TERMINOLOGICAL DEFINITION

intensional terminological definition (ITD): a TERMINOLOGICAL DEFINITION in which the CONCEPT being defined is shown to belong to a group of CONCEPTS (represented by the GENERIC CONCEPT, which is denoted by the term GENUS), and then distinguished from all other CONCEPTS within that group (i.e. from its CO–CONCEPTS) by virtue of its N+S CHARACTERISTICS (denoted by the term DIFFERENTIA). Also known as the Aristotelian definition, the traditional definition, and the classical definition.

ITD: see INTENSIONAL TERMINOLOGICAL DEFINITION (ITD).

KB: see KNOWLEDGE BASE (KB).

knowledge base (KB): in ARTIFICIAL INTELLIGENCE, a special type of database that contains a highly structured set of knowledge particular to a given SUBJECT FIELD. [LB]

knowledge engineering: a subfield of ARTIFICIAL INTELLIGENCE that is primarily concerned with acquiring, formalizing, and refining knowledge so that it can be easily used by machines and also by the people using the machines. [LB]

n+s characteristic: a NECESSARY CHARACTERISTIC which, combined with other NECESSARY CHARACTERISTICS, is sufficient to describe a CONCEPT. A set of CHARACTERISTICS {S} is sufficient for a CONCEPT (X), iff S implies X ($S \rightarrow X$); i.e. something that possesses the given set of n+s characteristics is necessarily the CONCEPT (X).

necessary and sufficient characteristic: see N+S CHARACTERISTIC

necessary characteristic: a CHARACTERISTIC possessed by all the OBJECTS denoted by a given TERM. A CHARACTERISTIC (C) is "necessary" to a CONCEPT (X) iff X must have/be C. For example, the CONCEPT **river** has the necessary characteristics ***origin:natural***, ***motion:flowing***, ***width:> W***, i.e. every river has these CHARACTERISTICS. Contrasts with TYPICAL CHARACTERISTIC.

object: any part of a perceivable or conceivable world. Note: objects may be material (e.g. engine) or immaterial (e.g. magnetism) [ISO 1990a:1].

partitive concept: a SUBCONCEPT in a partitive relation with a COMPREHENSIVE CONCEPT. For example, the CONCEPT **frame** is a partitive concept relative to the COMPREHENSIVE CONCEPT **bicycle**.

partitive relation: a HIERARCHICAL RELATION between a SUPERCONCEPT and SUBCONCEPT in which the SUPERCONCEPT refers to each object as a whole and each SUBCONCEPT to a part of it.

relation: a CHARACTERISTIC which provides information about other CONCEPTS that are related to a given CONCEPT as "peers," i.e. they are specialized CONCEPTS themselves (e.g. the CONCEPT **fax machine** could have the relations ***produces:fax***, or ***part:input tray***, in which information is provided about the CONCEPTS **fax** and **input tray**).

specific concept: a SUBCONCEPT in a GENERIC-SPECIFIC RELATION (i.e. it is more specialized than the GENERIC CONCEPT). A specific concept has all the CHARACTERISTICS of its GENERIC CONCEPT, plus at least one additional or modified CHARACTERISTIC, e.g. **compact disc** is a specific concept of **optical disc**. [LB]

subconcept: a CONCEPT in a hierarchical system which can be grouped together with at least one more CONCEPT of the same level to form a higher-ranking CONCEPT (referred to as a SUPERCONCEPT). [ISO 1990a:2] There are two types of subconcept: SPECIFIC CONCEPT and PARTITIVE CONCEPT.

subfield: a portion of a SUBJECT FIELD (e.g. **optical storage media** is a subfield of the **optical storage technology** subject field).

subject field: a section of human knowledge, delimted according to purpose.

subject field expert: a person who has studied the subject field intensively and who is recognized as a reliable and well-informed source of information about the subject field.

superconcept: a CONCEPT in a hierarchical system which can be subdivided into a number of lower-ranking CONCEPTS (referred to as SUBCONCEPTS). [ISO 1990a:1] Superconcepts can be of two types: GENERIC CONCEPTS and COMPREHENSIVE CONCEPTS.

template: a set of CHARACTERISTIC NAMES considered fundamental to understanding the high-level CONCEPTS of a given SUBJECT FIELD or SUBFIELD. The template is used to guide the terminologist's CONCEPT ANALYSIS work.

term: the designation of a specialized CONCEPT in a special language by a linguistic expression [ISO 1990a:5].

terminological definition: a DEFINITION in which a CONCEPT of a specialized SUBJECT FIELD is described and differentiated from other CONCEPTS within a CONCEPT SYSTEM.

terminological knowledge base (TKB): a hybrid between a conventional term bank and a KNOWLEDGE BASE as this is known in ARTIFICIAL INTELLIGENCE. [LB]

NOTE: This is the definition of terminological knowledge base used by the COGNITERM Project, as described in Meyer/Skuce/Bowker/Eck 1992a/b and Meyer/Bowker/Eck 1992a/b.

terminology: the discipline concerned with describing and naming CONCEPTS in specialized SUBJECT FIELDS.

TKB: see TERMINOLOGICAL KNOWLEDGE BASE (TKB).

typical characteristic: a CHARACTERISTIC not possessed by all the OBJECTS denoted by a given TERM, but nevertheless possessed by many of them (e.g. the CONCEPT **river** has the TYPICAL CHARACTERISTIC ***fauna:fish***: most rivers have fish, but some do not). Typical characteristics are not usually included in definitions. Contrasts with NECESSARY CHARACTERISTIC.

Introduction

The goal of this thesis is to develop and test a DEFINITION-ORIENTED APPROACH TO CONCEPT ANALYSIS within a knowledge-based framework. This approach, which allows the terminologist to construct INTENSIONAL TERMINOLOGICAL DEFINITIONS (ITDs) in a TERMINOLOGICAL KNOWLEDGE BASE (TKB), was developed in three main steps: (a) by exploring how a GENERAL APPROACH TO CONCEPT ANALYSIS (i.e. one that allows a terminologist to carry out CONCEPT ANALYSIS in a general way, without definition construction specifically in mind) is carried out using a generic knowledge management system called CODE (Conceptually Oriented Description Environment); (b) by gaining insights from the TERMINOLOGY literature, both on the nature of the ITD and on strategies for its creation;[1] and (c) by examining how certain features of CODE can be used to enhance the general approach to concept analysis so that ITDs can be constructed from the information that results from concept analysis in a systematic manner. Until now, terminologists have tended to use pencil-and-paper for much of their work; in this thesis, I argue that a knowledge-based computer tool such as CODE, used to integrate knowledge representation ideas from artificial intelligence with the traditional terminology approach, will help terminologists carry out this work in a more methodical and more conceptually-oriented manner.

Once I had developed the proposed approach, I tested it by creating a small TKB on lasers (a SUBJECT FIELD related to optical storage technology[2]), complete with ITDs. This test enabled me to refine the proposed approach, to recommend improvements to CODE, and to suggest possible directions for future research in knowledge-based definition construction. The final result is an

---

[1] The primary goal of this approach is to *construct* definitions, since a knowledge-based approach provides the "building blocks" for the definition (in the form of highly-structured conceptual characteristics). However, it could also be used to critique and/or modify existing definitions, which could then be added to the TKB.

[2] The subject field of lasers is related to optical storage technology in that lasers are a component of **optical storage devices** (e.g. **CD-ROM drives** and **CD players**), used to both write to and read from **optical storage media** (e.g. **CD-ROMs** and **CD-Audio discs**).

approach that terminologists can follow when carrying out DEFINITION-ORIENTED CONCEPT ANALYSIS using CODE, in which the aim is not only to gather information about a subject field's CONCEPTS and TERMS, but also to create ITDs from this information.

To understand the goal of this thesis, some explanation is required of both a knowledge-based approach to terminology (including the COGNITERM Project and knowledge-based concept analysis) and definitions in general. Each is discussed below.

0.1 A Knowledge-Based Approach to Terminology

Terminology can be described as the discipline concerned with describing and naming concepts in specialized subject fields, which implies that terminologists gather information both about the concept (conceptual information) and about the term(s) used to denote the concept (linguistic information). This information is presented in various forms by various means, including dictionaries, glossaries, and terminological databases (TDBs). Until recently, however, the focus has been more on linguistic information than on conceptual information, with the result being that extensive information about the concepts has not been made available to users. This deficiency has led an increasing number of terminology researchers to explore the possibility of TDBs that are conceptually richer than those of the previous generation [cf. Blampain/Petrussa/Van Campenhoudt 1992, Freibott/Heid 1990, Galinski 1988a/b, Meyer/Skuce/Bowker/Eck 1992a, Pearson 1992, Talbot/O'Croinin 1992]. In the view of some of these researchers, the key to such conceptual richness is a KNOWLEDGE BASE (KB) as it is known in artificial intelligence.

ARTIFICIAL INTELLIGENCE (AI) is a field in which terminology researchers have shown increasing interest. Of particular interest is a subfield known as KNOWLEDGE ENGINEERING, which is concerned with methods and technologies for building (or "engineering") knowledge bases. Some of the signs of this growing interest in a terminology-AI rapprochement include events such as the

International Congresses on Terminology and Knowledge Engineering, held in Trier in 1987 and 1990 and scheduled for Cologne in 1993; the formation of the International Association for Terminology and Knowledge Transfer; and the publication of such articles as "À la recherche d'écosystèmes terminologiques" [Blampain/Petrussa/Van Campenhoudt 1992], "Des fichiers terminologiques aux bases de connaissances" [de Bessé 1992] and "Towards a New Generation of Terminological Resources: An Experiment in Building a Terminological Knowledge Base" [Meyer/Skuce/Bowker/Eck 1992a].

Over the last two decades, knowledge engineering has emerged from the need to build computer systems which are geared towards solving problems in specialized fields of human expertise. These systems are thus knowledge-based: they must "know" what human experts know. Similarly, a growing number of terminology researchers feel that the TDBs of the future will also have to be knowledge-based and "know" what the experts know [Meyer/Paradis 1991:3].

Essentially, a knowledge base is a collection of information about the concepts of a specialized subject field. As well as being rich in content, the information is explicit and highly structured, making it relatively easy for the user (human or computer) to access it and learn from it. A TDB with this rich, explicit, highly-structured conceptual component has been called a terminological knowledge base, or TKB [cf. Meyer/Skuce/Bowker/Eck 1992a, Meyer/Bowker/Eck 1992a].

This development in terminology has been somewhat parallelled in that of lexicology, where researchers have been examining the usefulness of lexical knowledge bases, or LKBs [cf. Atkins 1991, Boguraev/Levin 1990, Pustejovsky/Bergler 1991]. While LKBs and TKBs have several similarities, there is one fundamental difference of particular interest to this thesis: while a LKB tries to capture what a native speaker knows about lexemes of general vocabulary, a TKB tries to capture what a native speaker *who is also a subject-field expert* knows about *specialized* lexemes (terms).

0.1.1 The COGNITERM Project

One project exploring a knowledge-based approach to terminology is the COGNITERM Project, being carried out at the Artificial Intelligence Laboratory at the University of Ottawa under the direction of Dr. Ingrid Meyer and Dr. Douglas Skuce. Its general objective is to investigate ways in which knowledge engineering and terminology can benefit each other. More specific objectives are (a) to develop an approach for carrying out concept analysis using a knowledge management system (CODE), and (b) to build a terminological knowledge base (COGNITERM) using the results of this concept analysis. Along the way, certain features of CODE are being developed to make it more suited to terminologists' needs.

I joined the COGNITERM Project at its inception in May 1991 as a research assistant. In the following months, the COGNITERM team developed a general approach to concept analysis, developed concurrently with the construction of a TKB for optical storage media (a SUBFIELD of optical storage technology).[3] This approach (explored in detail in part 1, chapter 3) contributes extensively to the definition-oriented approach to concept analysis developed in this thesis.

0.1.2 A Definition-Oriented Approach to Concept Analysis

In the past, definitions have been one of the few means[4] by which conceptual information has been furnished to TDB users. As definitions are an excellent means of providing users with a rapid overview of the most important information about a concept, it is desirable that they continue to be found within TKBs. For a variety of reasons, our early COGNITERM research was dominated by very basic methodological concerns, and constraints of time and resources precluded any specific

[3] This approach is documented in detail in the technical report "COGNITERM: Structure and Methodology" [Meyer/Skuce/Bowker/Eck 1992b]. More information about COGNITERM can be found in Meyer/Skuce 1990, Meyer 1991, Meyer 1992, Meyer/Skuce/Bowker/Eck 1992a and Meyer/Bowker/Eck 1992a.

[4] Other means have included subject-field indicators and contexts.

attention to definitions. Thus, our TKB was not designed specifically with a view to constructing definitions.

When it came time to start thinking about including definitions in the TKB, I realized that the general approach that we had been using would likely have to be modified to accommodate definition construction. But what exactly did this entail? How would we go about developing a definition-oriented approach? What distinguishes a definition-oriented approach from a general one? What steps of the general approach would have to be added, modified, or removed to accommodate definition construction?

These intriguing questions seemed like the perfect starting point for a Master's thesis. By finding the answers, I hoped to develop a definition-oriented approach to concept analysis, and at the same time contribute to the development of the COGNITERM Project and to knowledge-based terminology research in general.

0.2 Definitions In General

Before I could begin searching for any answers in particular, however, I had to find out more about definitions in general. My initial research revealed that there was much more to definitions than I had first suspected.

0.2.1 A Brief History of the Definition

Definitions were initially proposed for the study of philosophy nearly 2500 years ago, coming into prominence in the dialectics of Socrates and in Plato's attempt to define justice in the first book of *The Republic* [Miller 1980:38]. The concept of definitions was further developed by Aristotle, most notably in his *Topics* [Robinson 1965:1]. According to Lo Piparo [1990:24]:

> "Le concept de définition est en effet un concept central dans la théorie sémantique d'Aristote.... Il n'y a pas d'ouvrages d'Aristote où les problèmes du langage et de la

définition ne soient expressément discutés ou évoqués."

The fascination with definitions did not disappear with the ancient Greeks. Marcus Aurelius, the Roman emperor, was also acutely aware of their importance, and wrote [Borsodi 1967:3]:

> Make for thyself a definition or description of the thing which is presented to thee, so as to see distinctly what kind of thing it is...and tell thyself its proper name, and the names of the things of which it is compounded, and unto which it will be resolved. For nothing is so productive of elevation of mind as to be able to examine methodically and truly every object which is presented to thee in life, and always to look at things so as to see at the same time what kind of universe this is, and what kind of use everything has with reference to the whole, and what with reference to man.

In the seventeenth century, Thomas Hobbes, the philosopher, also stressed the need for definitions in his *Leviathan* [Borsodi 1967:3], writing:

> There can be nothing so absurd but may be found in the books of philosophers. And the reason is manifest. For there is not one of them that begins his ratiocination from the definition, or explications of the words they are to use; which is a method that hath been used only in geometry, whose conclusions have thereby been made indisputable.

Despite thousands of years of examination, however, definitions continue to be a topic of study. As Borsodi [1967:4] notes, "thoughtful men are still calling attention to the problem [of representing knowledge], because obviously nothing has been done to dispose of it."

0.2.2 Defining the Definition

Many of those with an interest in the subject have attempted to answer the question "What is the definition of *definition*?" Some of the most famous answers are summarized in Robinson [1965:2]:

1. Plato (philosopher):
   (a) the revelation of one's thoughts by means of speech;
   (b) the ability to answer someone by means of the elements of the thing, when he asks you what a thing is;
   (c) the ability to give some mark by which the thing asked about differs from all things [*Theaetetus,* 206C–7A, 208C].

2. Aristotle (philosopher): the account of the essence of things [*Topics,* I 5].

3. Cicero (politician and orator): a certain brief and circumscribed account of the properties of the thing we wish to define [*De Oratore,* I 42, 189].

4. John Milton (poet and essayist): that which refines the pure essence of things from the circumstance [*Works* (1851), IV 168].

5. Baruch Spinoza (philosopher): the true definition of each thing involves nothing and expresses nothing but the nature of the thing defined [*Ethics,* I, prop. 8, n.2].

6. John Locke (philosopher): the ability to make another understand by Words, what Idea the Term defin'd stands for [*Essay,* III.iii.10].

7. Emmanuel Kant (philosopher): the presentation of the complete, original concept of a thing within the limits of its **concept** [*Critique of Pure Reason,* A 727, trans. N.K. Smith].

8. John Stuart Mill (philosopher and economist): a proposition declaratory of the meaning of a word; namely, either the meaning which it bears in common acceptation, or that which the speaker or writer, for the particular purposes of his discourse, intends to annex to it.

9. Alfred North Whitehead and Bertrand Russell (both: mathematician, logician, philosopher): a declaration that a certain newly introduced symbol...is to mean the same as a certain other combination of symbols of which the meaning is already known [*Principia Mathematica,* 2nd ed., 11].

10. Ludwig Josef Wittgenstein (logician, philosopher): rules for the translation of one language into another [*Tractatus,* 3. 343].

11. Rudolf Carnap (logician, philosopher): a rule for mutual transformation of words in the same language [*The Unity of Science,* trans. Max Black, 39].

For the purposes of this thesis, I will define definition as "a statement about a concept which provides the essence of the concept, thus distinguishing it from related concepts."

The list above presents just some of the better known views on definitions; in fact, there is an almost overwhelming amount of material available. Not only have whole chapters been written on the subject (cf. Chapter 2: "Definition" in Miller 1980, "La définition terminologique" and "Terminological Definitions: Characteristics and Demands" in Duquet-Picard 1982, "Definition" in Landau 1984), but entire books have been devoted to it (cf. Chaurand 1990, Robinson 1965, Borsodi 1967). There have even been conferences, including the *Colloque la Définition* (Paris 1988), dedicated to exploring the definition.

Needless to say, this abundance of information reflects the profusion of opinions on what a definition is.

0.2.3 Definition Typologies

Along with discussing what constitutes a definition, many experts also provide some kind of definition typology. Predictably, there is little agreement on definition types. For example, Borsodi [1967:18] feels that are at least 25 distinct types of definition, which he classifies as either verbal or nonverbal:

DEFINITIONS			
Verbal			Nonverbal
Derivative	Operational	Quotational	Ocular
Translational	Descriptive	Analogic	Pictorial
Recommended	Historical	Obsolescent	Aural
Synonymic	Anatomic	Archaic	Tactile
Antonymic	Qualitative	Binomial	Palatal
Meronymic	Quantitative		Olfactory
Classificatory	Illustrative		

Robinson [1965:7], on the other hand, lists some of the definition types proposed over the past two centuries:[5]

1. Operational definition (P.W. Bridgman 1928)
2. Genetic definition (J.E. Creighton)
3. Implicit definition (J.D. Gergonne 1818)
4. Ostensive definition (W.E. Johnson 1921)
5. Analytical definition (W.E. Johnson 1921)
6. Synthetic definition (W.E. Johnson 1921)
7. Extensive definition (J.N. Keynes)
8. Successive definition (V.F. Lenzen 1938)
9. Denotative definition (A.W. Levi/A.M. Frye 1941)

[5] Further information about the author and work have been provided by Nkwenti-Azeh [1989:162], who noted that, in the cases where no dates are given, the author and/or work referred to by Robinson could not be identified.

10. Connotative definition (A.W. Levi/A.M. Frye 1941)
11. Definition by description (C.I. Lewis 1946)
12. Equational definition (S.C. Pepper 1945)
13. Descriptive definition (S.C. Pepper 1945)
14. Co-ordinating definition (H. Reichenbach 1920)
15. Persuasive definition (C.L. Stevenson 1944)
16. Definition in use (Whitehead/Russell 1925)

As Nkwenti-Azeh [1989:162] notes (in a Ph.D. thesis on TERMINOLOGICAL DEFINITIONS), Robinson's list of definitions does not include the Aristotelian categories of definition (i.e. nominal vs. real), nor does it take into account the three so-called "purposes" of definition which Robinson himself proposes: word-word definition, word-thing definition, and thing-thing definition.

To Robinson's list can be added definitions that Nkwenti-Azeh [1989:162] has gathered from linguistic and philosophical publications that have appeared in the last two decades:

17. Complete definition (Bierwisch/Kiefer 1969)
18. Partial definition (Bierwisch/Kiefer 1969)
19. Proper definition (Bierwisch/Kiefer 1969)
20. Redundant definition (Bierwisch/Kiefer 1969)
21. Generic definition (Dahlberg 1979)
22. Partitive definition (Dahlberg 1979)
23. Functional definition (Dahlberg 1979)
24. Definition by opposition (Dahlberg 1979)
25. Conceptual (or real) definition (Dahlberg 1979)
26. Functional definition (Dahlberg 1979)
27. Substitutive definition (Frawley 1982)
28. Reductive definition (Frawley 1982)
29. Synthetic definition (Frawley 1982)
30. Metaphorical definition (Frawley 1982)
31. Syntactic definition (Frawley 1982)
32. Pragmatic definition (Frawley 1982)
33. Definition by intension (Felber 1983)
34. Definition by extension (Felber 1983)
35. Definition by context (Felber 1983)

This rather large typology results from different authors using different criteria to classify definitions. Although a comparison of the different types reveals some synonymy, there are still many from which to choose. Despite all the disagreement, however, there appears to be some consensus that the *intensional* definition is particularly important; it is by far the most common type of definition, mentioned in virtually all works discussing definitions.

0.3 The Focus of This Thesis

Given the huge body of literature on definitions, which has produced a correspondingly wide range of definition types, I have restricted my exploration of the definition, as noted below.

0.3.1 The Intensional Terminological Definition (ITD)

Obviously, a terminological framework presupposes that the focus is on terminological definitions,[6] i.e. the definitions of concepts found in specialized subject fields. Ever since Eugen Wüster founded the discipline of terminology some 60 years ago, terminologists, with their mandate to explore and analyze the terms and concepts of specialized subject fields, have been interested in these definitions "for practical purposes" [Dahlberg 1983:15].

While there have been varying opinions on the importance of definitions within terminology, the current consensus appears to be that the definition is indeed a useful tool,[7] and indeed, one finds them in most term banks and in many specialized dictionaries. Cole [1987:85] points out one of the main reasons for this view:

> Although the nature of a terminologically adequate definition is still the subject of active debate..., it has been recognized that the requirement for completeness of information about a given concept can best be met in this way.

---

[6] While I recognize that the differences between lexical (or general) definitions and terminological definitions have been the subject of great debate, a comparison of them is beyond the scope of this thesis. Nonetheless, some of the basic literature on lexical definitions was consulted, including Brustkern/Hess 1982, Hartmann 1983, and Landau 1984.

[7] There appears to be agreement that definitions are useful and important; however, not all researchers consider a definition to be *the most important* component of a term bank entry. For example, Harris [1983:143] contends that translators are primarily interested in the foreign language equivalents found in term banks, and "provided they have an entry word and a translation, translators are prepared, if necessary, to forego anything else." Skaikevitch/Oubine [1988:13] surveyed translators and researchers about bilingual terminological dictionaries and reported that "...49 persons [out of 224 persons polled] were in favour of inclusion of definitions in the dictionary entry in all cases. The majority (175) think that such inclusion is justified only in special cases.... 13 respondents were resolutely against the inclusion of definitions in specialized dictionaries."

Rousseau [1983:38] adds:

> La raison d'être de la terminologie est avant tout de savoir nommer les choses, dans une perspective de connaissance (aspect cognitif) et d'expression ou de communication. Et pour pouvoir utiliser une terminologie adéquatement, c'est-à-dire avec précision et efficacité, il faut définir.

I have restricted my study to the intensional terminological definition, also called the Aristotelian or classical definition. The ITD can be described as a terminological definition in which the CONCEPT being defined is shown as belonging to a group of concepts (represented by the GENERIC CONCEPT) and then distinguished from all other concepts within that group by virtue of its CHARACTERISTICS. It is the oldest and most common type of definition, devised by Aristotle. It is well suited to terminology, as it appears to be particularly appropriate for defining nouns [Miller 1990:246], which make up "approximately 80% of the terminology of a special field" [Baudot/Clas 1984:49, as cited in Béjoint 1988:356]. As Dubuc [1985:96] notes of the intensional definition, "Ce mode de définition, largement utilisé en lexicographie, sert aussi abondamment en terminologie."

There are other interesting types of terminological definition, including the extensional, partitive, and functional, each having its own features (see Picht/Draskau 1985 and Sager 1990 for overviews). These features must be explored in detail to be fully understood, however, and such an exploration cannot be done within the scope of this thesis. A definition-oriented approach to concept analysis developed for ITDs should, however, provide a useful starting point for further research on the construction of other types of terminological definition.

0.3.2 The Method of Defining

While much of terminology literature discusses selecting definitions from contexts or modifying contextual definitions as the method of defining, I have chosen the third option, hand-crafting definitions. It should be noted, however, that the work carried out within this thesis has implications for the other two methods: it could be assumed that if one knows how to hand-craft definitions, one is also qualified to select or modify them.

0.3.3 The Subject Field

This thesis is also restricted to two subject fields: optical storage technology for the development of the definition-oriented approach to concept analysis, and lasers for the subsequent testing of this approach. Both fields seemed appropriate for exploration, because (a) I had some prior knowledge of them as a COGNITERM researcher; (b) they are both technical (it is often said that concepts in scientific/technical subject fields are easier to define than those in the so-called "soft sciences"; and (c) I had access to local subject-field experts, a fact which became increasingly important as the thesis advanced.

Clearly, I cannot claim that the approach developed will work outside the confines of these particular subject fields or, particularly, outside scientific/technical ones. Testing them on other subfields would make interesting material for future research, and I hope the present work provides a useful starting point.

0.3.4 The Language

While COGNITERM as a whole is being developed as a bilingual (French-English) TKB, the small TKB on lasers that I created to test the definition-oriented approach is unilingually English, both for simplicity and because laser technology has been developed primarily in the English-speaking world (with the result that there is little difference between concepts in the two languages). It should be noted, however, that bilingual issues are discussed regularly throughout this thesis. I would like to see this thesis used as a springboard for future work on the bilingual aspect of ITDs created in a knowledge-based environment, particularly for subject fields which show significant conceptual differences between languages.

0.3.5 The COGNITERM Framework

Since I am working within the COGNITERM framework (with concept analysis carried out using CODE and the resulting information placed in a TKB), the definition-oriented approach to concept analysis is proposed within that context. Therefore, this thesis makes no claims that the proposed approach will work within other technological environments. However, it should be noted that CODE has been designed as a generic tool, and thus can be used in many other applications,[8] unlike other knowledge engineering tools that have been designed for one application only. Since the proposed approach has been developed for a generic environment, it may eventually be adapted for other knowledge engineering tools as well.

0.4 The Structure and Methodology of the Thesis

This thesis is divided into three parts. Part 1 describes the general research framework being used. This includes a discussion of (a) the shared basic concepts of knowledge engineering and terminology, (b) the technology being used (i.e. the CODE system), and (c) the general approach to concept analysis developed within the COGNITERM Project.

Part 2 proposes the initial definition-oriented approach to concept analysis, which is derived from (a) an exploration of what is said in the terminology literature (and to some degree, the lexicology literature) about intensional definitions, as a product, a process, and a tool; (b) my participation in the development of one particular CODE feature called the Characteristic Comparison Matrix, which facilitates the construction of definitions; and (c) my own insight, based on practical experience as a COGNITERM researcher, into how the general approach to concept analysis could be enhanced to accommodate the construction of definitions. The resulting approach consists of a series

[8] Some other applications of CODE that have been explored (besides terminology) include software engineering, software documentation, and data management.

of non-sequential steps which, if followed by the terminologist, should result in good quality, consistent ITDs.

Part 3 describes the testing of the proposed approach, with the result being a small TKB on the subfield of lasers, complete with conceptual information, linguistic information, and ITDs. This is followed by an assessment of the test results, including suggestions for refinements to the initial definition-oriented approach to concept analysis, improvements in CODE, and other possible directions for future research.

By exploring the development and testing of a definition-oriented approach to concept analysis within a TKB, this thesis represents an important step in the ongoing development of a knowledge-based approach to terminology. I hope that this approach will guide terminologists in the systematic construction of ITDs within a TKB, and provide knowledge-based terminology researchers with a better understanding of the methodological and technological issues that the age-old problem of definition construction continues to raise.

PART 1

GENERAL CONCEPT ANALYSIS: BASIC CONCEPTS, TECHNOLOGY, AND APPROACH

The approach that I have developed for carrying out concept analysis specifically aimed at constructing definitions is termed a *definition-oriented approach to concept analysis*. It has been developed by enhancing the general approach to concept analysis that was developed by the COGNITERM team. If this definition-oriented approach is to be fully understood, several things about its forerunner must first be explained. Part 1 of this thesis, divided into three chapters, provides the necessary background.

Chapter 1 explains the basic concepts behind the general approach (concepts common to both knowledge engineering and terminology). Chapter 2 describes the technology being used to help develop this approach (the CODE system). Chapter 3 presents the general approach that has been developed by COGNITERM researchers to help the terminologist create a terminological knowledge base within a knowledge-based framework.

Chapter 1 THE BASIC CONCEPTS: CONCEPTS, CHARACTERISTICS, AND RELATIONS

1.0 INTRODUCTION

Much of terminology work involves concept analysis, which entails describing the characteristics of individual concepts and the RELATIONS that hold within CONCEPT SYSTEMS. According to Skuce/Meyer [1990a:1], who work within a knowledge-based framework:

> It [concept analysis] is the central concern of terminology, essential to carrying out the knowledge acquisition that always precedes and accompanies the linguistic work, delimiting and partitioning subject fields, constructing definitions, distinguishing quasi-synonyms, dealing with neology, and communicating with subject-field experts and colleagues.

This opinion has also been espoused by more traditional terminologists, such as Dubuc [1985:20], who writes that "...l'identification de la notion [the concept] par l'analyse contextuelle apparaît comme la méthode fondamentale de la recherche terminologique."

It seems logical, then, before discussing any aspect of terminology (whether knowledge-based or traditional), to examine more closely the basic concepts underlying concept analysis, namely concepts, characteristics, and relations. These three basic concepts are closely linked: concepts are described in terms of their characteristics, while relations are a particularly important kind of characteristic, since they contribute to the concept network.

This chapter examines concepts, characteristics , and relations more closely. It also describes how concept analysis can help in the practical activities of terminology outlined in the quotation above.

1.1 THE CONCEPT

The concept is ranked highly within the discipline of terminology, as "all modern work on the problems of terminology takes as its starting point the *concept* itself" [Wüster 1968:xii].

1.1.1 Defining the Concept

Although there is general consensus within the terminology community that the concept is important, there is less agreement on what exactly a concept is. Sager [1990:23] provides a selection of definitions formulated over the years by various standardization committees, along with some definitions of his own:

— Concepts are mental constructs, abstractions which may be used in classifying the individual OBJECTS of the inner and outer world [British Standard Recommendation for the selection, formation and definition of technical terms, BS.3669:1963].

— The objects of all fields of knowledge and human activity, such as things, their properties, qualities, phenomena, etc., are represented by concepts [UK proposal for the revision of ISO document R 704].

— A concept is a mental construct for classifying individual objects of the outer and inner world by means of a more or less arbitrary abstraction [1968 draft version of ISO Standard 704].

— A concept is a unit of thought, produced by grouping individual objects related by common characteristics [draft of a German DIN document].

— A concept is a coherent group of judgements concerning an object whose nucleus is made up of those judgements which reflect the inherent characteristics of the object. [Soviet Union proposal for the revision of ISO document R 704].

— A concept is any unit of thought [Final version of the Draft International Standard ISO/DIS 704, 1985].

— A concept [is] a unit of knowledge [Sager 1990:15].

— [A concept is a] construct of human cognition processes which assists in the classification of objects by way of systemic or arbitrary abstraction (provisional definition [Sager 1990:22]).

Finally, the *International Standard ISO 1087 Terminology—Vocabulary* provides the following definition, to be considered the definition of concept for this thesis:

— A concept is a "unit of thought constituted through abstraction on the basis of properties common to a set of objects."[9] [ISO 1990a:1]

[9] ISO [1990a:1] describes an object as "any part of the perceivable or conceivable world," and notes that objects may be material (e.g. engine) or immaterial (e.g. magnetism).

There is an obvious divergence of opinion on what exactly a concept is. Sager [1990:22] himself suggests that the "concept be considered another axiomatic primitive, like 'word' or 'sentence,' [and] conveniently left undefined." Picht/Draskau [1985:36] sum up the situation by noting that "if we are looking for an answer to the question 'what is the concept?' we may expect to receive very different answers according to the branch of learning and the school," and that it is "highly questionable whether it will ever be possible to achieve unanimous agreement on the nature of the concept; it does not seem likely."

1.1.2 Types of Concept

Concept typologies are found in the terminology literature; however, there is no one generally accepted typology. Sager [1990:26–27], for example, groups concepts together on a number of levels:

— entities (derived by abstraction from material or abstract objects)

— activities (processes, operations, and actions, performed with, by, or for entities)

— qualities (properties used to differentiate entities)

— relations (established between any of the other three types of concepts)

Picht/Draskau [1985:38–39], however, divide concepts into general concepts and individual concepts. A general concept represents an "indefinite" object, e.g. cathedral, and may be linguistically realized by a term. An individual concept represents a "definite" or unique object, e.g. Cologne Cathedral, linguistically realized by a name.

Picht/Draskau's concept typology corresponds to one found in knowledge engineering; Skuce [1991b:5], for example, distinguishes two main kinds of concept: type concepts (e.g. city) and instance concepts (e.g. Paris). It should be noted, however, that terminology concentrates on general concepts (as does knowledge engineering).

Concepts can also be classified according to their position in a hierarchy. This classification is described in some detail, since the intensional definition is based on the generic–specific HIERARCHICAL RELATION. ISO [1990a:1–2] defines several types of concept found in a hierarchy:

I. superordinate concept (in this thesis, SUPERCONCEPT):[10] concept in a hierarchical system which can be subdivided into a number of lower-ranking concepts

   A. generic concept: superordinate concept in a generic relation (i.e. more general than the SPECIFIC CONCEPT)

   B. COMPREHENSIVE CONCEPT: superordinate concept in a partitive relation[11] (i.e. a concept which, as a whole, is divided into parts)

II. subordinate concept (in this thesis, SUBCONCEPT): concept in a hierarchical system which can be grouped together with at least one more concept of the same level to form a higher ranking concept

   A. specific concept: subordinate concept in a generic relation (i.e. a more specialized version of the generic concept)

   B. PARTITIVE CONCEPT: subordinate concept in a partitive relation (i.e. a concept which is part of a whole)

III. coordinate concept (in this thesis, CO-CONCEPT): concept in a hierarchical system (whether generic–specific or part–whole[12]) which ranks at the same level as one or more other concepts

For the purposes of this thesis, the following types of concept will be considered:

— from Sager's list: entities

— from Picht/Draskau's list (and according to the primary thrust of terminology work): general concepts

— from the ISO list: all types, although much less emphasis is placed on comprehensive and partitive concepts

[10] Although the ISO 1087 Standard [ISO 1990a:1–2] uses the terms *superordinate concept, subordinate concept,* and coordinate concept, I use the less unwieldy terms *superconcept, subconcept,* and *co–concept* because they are accepted terms in the field of AI, which is contributing more and more to computational terminology.

[11] For reasons beyond the scope of this thesis, it is COGNITERM convention to use the more scientific term "meronymic relation" rather than partitive relation. Therefore, this term will be seen on figures throughout this thesis.

[12] While the ISO 1087 Standard [ISO 1990a] does not provide any terms to distinguish generic–specific coordinates from part–whole ones, Cruse [1986:137] refers to "sister concepts" in a generic–specific hierarchy (or taxonomy, in his terms) as co–taxonyms; in a part–whole hierarchy (which he calls a meronomy) they are co–meronyms [Cruse 1986:160].

Although there appears to be little consensus on what a concept is, or on its typology, there does seem to be agreement that characteristics are used to describe a concept [cf. Sager 1990:23, Picht/ Draskau 1985:40, Boutin-Quesnel/Bélanger/Kerpan/Rousseau 1985:18, Cole 1987:78]. These conceptual "building blocks" are discussed further in section 1.2.

1.1.3 Denoting the Concept

While the concept may be considered the principal focus of terminological work, the term is extremely important for denoting the concept. The ISO 1087 Standard [1990a:5] defines a term as the "designation of a...concept in a special language by a linguistic expression."

ISO [1990b:3] makes a special point of noting that *term* is not synonymous with *concept*; term is the "symbol" [Sager 1990:22] used to represent the concept. This may seem confusing to laypersons at first, especially when the most obvious way to discuss concepts is via their terms. However, all one has to do is denote one concept in more than one language, and the difference becomes clear. The "thing that you play in your CD player" is denoted by the term *compact disc* in English and by *disque compact* in French. Regardless of the term used, however, the concept being denoted in each case is the same.[13]

1.2 CONCEPTUAL CHARACTERISTICS

As noted above, characteristics describe concepts. The ISO 1087 Standard [ISO 1990a:2] defines a characteristic as a "mental representation of a property of an object, serving to form and delimit its concept." Picht/Draskau [1985:40] note that the "sum of predications (=characteristics) is

[13] Concepts may vary from language to language; as Miller [1993:146] notes, "from the point of view of certain theorists, concepts are tied to particular languages." However, these issues are beyond the scope of this thesis, which deals with highly technical concepts that do not display obvious differences between English and French. "Fortunately, in technical domains (i.e. languages for special purposes...), there is a greater degree of correspondence between languages because of the international scope of most scientific activity and extensive efforts to standardize technical terminology" Miller [1993:147].

equal to the sum of our knowledge about a concept." The greater the number of characteristics identified for a given concept, the more is known about that concept.

The sum of characteristics that constitute a concept is its INTENSION, while the range of objects to which a concept refers is its EXTENSION [Sager 1990:24, Picht/Draskau 1985:40]. A concept identified by many characteristics is said to have a broad intension; consequently, its extension is narrow, as it can refer to only a limited number of objects. A concept distinguished by few characteristics has a narrow intension; since many objects can be included in its scope of meaning, it is said to have a broad extension.

1.2.1 Characteristic Components

Although the characteristic is often discussed as though it were a single entity, it can in fact be considered as having two components: the characteristic name (CN) and the characteristic value (CV). When a characteristic is mentioned in the terminology literature, it is usually its value which is being discussed. For example, Picht/Draskau [1987:47] provide a sequence of "characteristics" for the concept **table**:

(a) wooden

(b) round

(c) 70 cm high

The characteristic values are provided here, leaving the reader to infer the characteristic names (perhaps in this case: (a) material, (b) surface shape, and (c) height). Ideally, the characteristic should be explicitly expressed, with both the name and value specified. This explicitness has a twofold benefit. First, it clearly explains the characteristic to the user (i.e. it doesn't force the user to infer what is being discussed). Second, it helps ensure that duplicate, contradictory, or improper characteristics (i.e. values) are not included with the information gathered about a concept.

For the purposes of this thesis, the characteristic will be always be displayed in its full form:

characteristic name:*characteristic value (CN:CV).*

Naturally, if two or more concepts have the same CN:CV, it can be said that they have the same characteristic. However, if the CN is the same for two or more concepts, but the CVs are different, then the characteristic should be discussed in terms of its value; the CV can be said to be different, or in the case of a superconcept and subconcept, the CV can be said to have been modified.

1.2.2 Characteristics as Concepts

Picht/Draskau [1985:40] note that "characteristics are in themselves concepts," which can be taken to mean that they consider the CV to be a specialized concept. This statement should not be taken at face value, however: it is important to keep in mind that some characteristic values denote specialized concepts, while others denote general ones. Specialized concepts are found within a particular subject field; "unspecialized" concepts belong to general world knowledge. It is the specialized concepts which are of interest here.

For example, Table 1 displays some of the characteristics found for **CD–ROM**, including the characteristic ***error correction code:****LECC (layered error correction code)*. According to the subject-field literature, the CV of this characteristic, *LECC*, is a specialized concept. It has thus been included elsewhere in the COGNITERM optical storage technology TKB. The CV of the characteristic ***diameter:****5.25 inches*, however, is obviously not a specialized concept, and thus has not been placed in the TKB.

It is also possible for the characteristic name to be considered a specialized concept. The CN of ***error correction code:LECC***, seen in Table 1, happens to be a specialized concept, and is included in the optical storage technology TKB. The CN of ***diameter:5.25 inches***, on the other hand, can be considered an unspecialized concept; it is thus not included in the TKB.

SOME CHARACTERISTICS OF CD-ROM	
Characteristic Name	*Characteristic Value*
degree of writability	*read-only*
encoding method	*digital*
error correction code	*LECC (layered error correction code)*
physical standard	*Yellow Book*
diameter	*5.25 inches*

Table 1. Characteristics of **CD-ROM**: name and value.

In this thesis, a CV's concept status (specialized or unspecialized) is important for the classification of characteristics; this is discussed further in section 1.2.3.2.1.

1.2.3 Types of Characteristic

There are many different characteristic typologies, according to the literature. These typologies, some of which are described below, have inspired the two characteristic typologies used in this thesis. These two typologies are to be considered complementary, rather than opposing.

1.2.3.1 Types of Characteristic in the Literature

As Picht/Draskau [1987:44] note, no uniform classification of conceptual characteristics exists, despite considerable efforts. Picht/Draskau [1987:47–48] suggest some typologies, however, including:

— dependent vs. independent characteristics: dependent characteristics must exist in the concept system before more specialized concepts can be added, while independent characteristics do not. The example provided is that of the concept system **agriculture machine — harvester — combine harvester** (going from generic to specific). According to Picht/Draskau, the characteristic *harvest*[14] must already exist in the system (i.e. describing the concept **harvester**) before it is possible to proceed to the next level, where concepts such as **combine harvester**, with the characteristic *special type of harvest*, could be found.

— ordering characteristics vs. insignificant characteristics:[15] ordering characteristics are used to classify concepts, while insignificant characteristics are not; for example, for botanists classifying roses, ***shape of flower*** could be an ordering characteristic, while ***colour*** would be an insignificant characteristic.

— simple characteristics (e.g. *round*) vs. composite characteristics (e.g. *marxist–leninist*).

Sager [1990:24] groups characteristics into two main categories:

— essential, or NECESSARY AND SUFFICIENT, characteristics: those characteristics necessary to sufficiently distinguish a given concept from any other closely related concept (e.g. for **table:** *flat surface*, *at least one leg*);

— inessential characteristics: those characteristics which are observable in the individual object (e.g. for a given table: its colour and material)

Dubuc [1985:21] discusses characteristics briefly, referring to them as "éléments significatifs, composants de la notion." He provides a list of what he feels are the "traits sémantiques importants" of the concept, which include its "nature, fin, fonction, matière..." [1985:41]. Characteristics that are

[14] Although Picht/Draskau only provide what appears to be the characteristic value, one possible characteristic name could be ***function***. The whole characteristic is thus ***function:harvest***.

[15] Picht/Draskau note that this is considered to be a variable classification, since a characteristic may change categories through a change in point of view. For example, *colour* could be an ordering characteristic for horticulturists.

specifically found in definitions are termed by Dubuc as "caractéristiques essentielles" [1985:95].

While these various typologies differ, they do indicate that there is a "continuum of importance" for characteristics, with what I will call necessary characteristics (all instances of that concept having the characteristic) at one end, and typical characteristics (not all instances of that concept having the characteristic) at the other.

Within the field of artificial intelligence, there has been prolonged discussion but little agreement on characteristic types. However, there does seem to be general agreement on two major categories [cf. Skuce 1991a/b, Nirenburg *et al.* 1988], namely:

— relations: characteristics which provide information about other concepts related to a given concept as "peers," i.e. they are specialized concepts themselves (e.g. the concept **fax machine** could have the relations ***produces:fax***, or ***part:input tray***).

— ATTRIBUTES: characteristics which are inherent to the concept being described, not involving any relation to other specialized concepts (e.g. the concept **fax machine** could have the attributes ***height:30 cm, width:40 cm, weight:20 kg***).

Here, the typology is not based on the characteristic's "degree of importance," but rather on the relationship between the characteristic and the concept describes.

1.2.3.2 Characteristic Distinctions Within this Thesis

There are two complementary ways of classifying characteristics, each of which draws from the categories discussed above. The first distinction is made between relations and attributes; it can be used to classify characteristics when they are initially added to a TKB. The second distinction, which distinguishes characteristics according to importance, is used to classify them according to their usefulness in definition construction.

Distinction 1: relations vs. attributes

According to this distinction, a characteristic can be either a relation or an attribute, as seen below in Figure 1.

I. **CHARACTERISTIC:** A unit of information which helps to describe a concept.

A. **RELATION:** A characteristic which links a given concept (X) with another specialized concept (Y) (i.e. Y is the value of a characteristic of X).

1. **Hierarchical relation:** A relation involving a superconcept and a subconcept.

a. ***generic-specific relation:*** A hierarchical relation involving a generic concept and a specific concept; generic-specific relations allow characteristics to inherit from the generic concept to the specific concept.

b. ***partitive relation:*** A hierarchical relation involving a comprehensive concept and a partitive concept; partitive relations do not normally allow characteristics to inherit from the comprehensive concept to the partitive concept.

2. **Associative relation:** A non-hierarchical relation (non-hierarchical relations do not allow characteristics to inherit from one concept to another).

B. **ATTRIBUTE:** A characteristic which is inherent to the concept being described. Its value does not involve any other concept.

Figure 1. Characteristic Distinction 1: Relations vs. Attributes.

It may be surprising to see relations considered as characteristics, but if characteristics are thought of as "units of information which describe a concept," this conclusion is logical. A characteristic provides information about a concept. The relation that a given concept (X) has with another concept constitutes information about X. Therefore, a relation can be considered a characteristic. Each of these characteristic types is further discussed below.

A. Relations

The relations that exist between a given concept and other concepts are important for describing both the concept and the resulting concept systems. As Sowa [1991:179] notes:

> None of these words [sin, carburetor, tax shelter] can be understood in isolation.... The entry for *sin*, for example, might define it as a transgression against God. But that introduces the concepts of transgression and God. A transgression is a violation of a law, but that raises questions about how God gives laws and how they differ from human laws or laws of physics. A few more steps lead to the concepts of heaven and hell and eventually all of theology.... In every field of human endeavor, from cooking and fashion to topology and quantum mechanics, the basic concepts can only be understood in relation to other concepts in tightly organized structures of thought. Knowledge acquisition may begin with words, but it must also find the connections that link those words in larger structures.

ISO [1990a:3] recognizes two main types of relation: hierarchical relations and non-hierarchical (or associative) relations. As hierarchical relations are more relevant to definition construction, and have also been the object of much more research than associative relations, they will be discussed in greater detail.

1. Hierarchical relations

A hierarchical relation is a concept relation that is marked by the progressive division of a concept into narrower concepts called subconcepts or, conversely, by the progressive combination of concepts into broader concepts called superconcepts. There can be one or more levels of subordination or superordination, depending on the number of divisions. As discussed in section 1.1.2, concepts at the same level of subordination or superordination are called co-concepts. It should be noted that a subconcept in one relation may be considered a superconcept in another. Figure 1 displays the co-concepts **CD-ROM, CD-Audio,** and **CD-Video** in a generic-specific hierarchy.

ISO [1990a:3], consistent with most of the terminology literature (e.g. Picht/Draskau 1985, Sager 1990), recognizes two main types of hierarchical relation: GENERIC-SPECIFIC RELATIONS and PARTITIVE RELATIONS.

a. Generic-specific relations

According to ten Pas [1991:80], the generic-specific relation "is the most commonly used to structure a subject field conceptually." It is a particularly powerful relation, since it allows INHERITANCE of information (captured in characteristics) from superconcept to subconcept. In a generic-specific relation, there is a partial correspondence between the intensions of the generic concept and the specific concept. The specific concept inherits all of the characteristics of the generic concept, but differs from it by at least one modified or additional characteristic.

Figure 2 shows how the CODE graph designates the generic-specific relation between two concepts with an 's'. The concept of inheritance is discussed further in section 1.2.4.

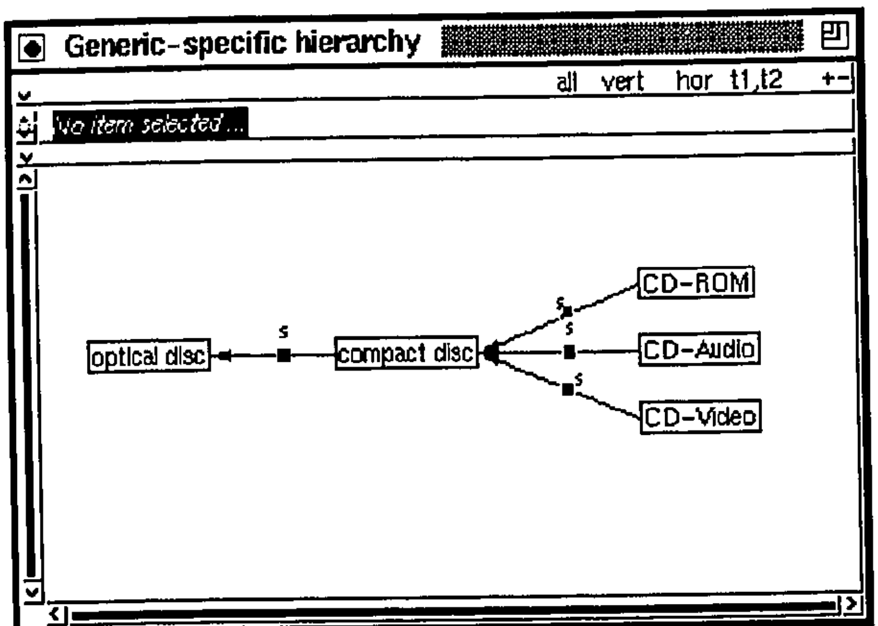


Figure 2. Generic-specific relations.

The generic-specific relation can be thought of as an "isa" or "kind of" relation, as it is usually expressed as "X is a Y," "Xs are Ys," "X is a kind of Y," "Xs are a type of Y." Figure 2 shows a series of generic-specific relations: **CD-Audio**, **CD-ROM**, and **CD-Video** are all kinds of **compact disc**, which is a kind of **optical disc**. Since the concepts **CD-Audio**, **CD-ROM**, and **CD-Video** share the same generic concept, they are co-concepts.

b. Partitive relations

In partitive relations, the comprehensive concept represents the object as a whole, and partitive concepts represent parts of this whole. According to Miller [1990:255], partitive relations tend to occur most frequently with concepts that represent physical objects. Richard [1977:91] concurs, noting: "...ces rapports sont évidemment les plus appropriés à la définition des objets du monde concret, si nombreux en technique." An example of partitive relations can be found in Figure 3.

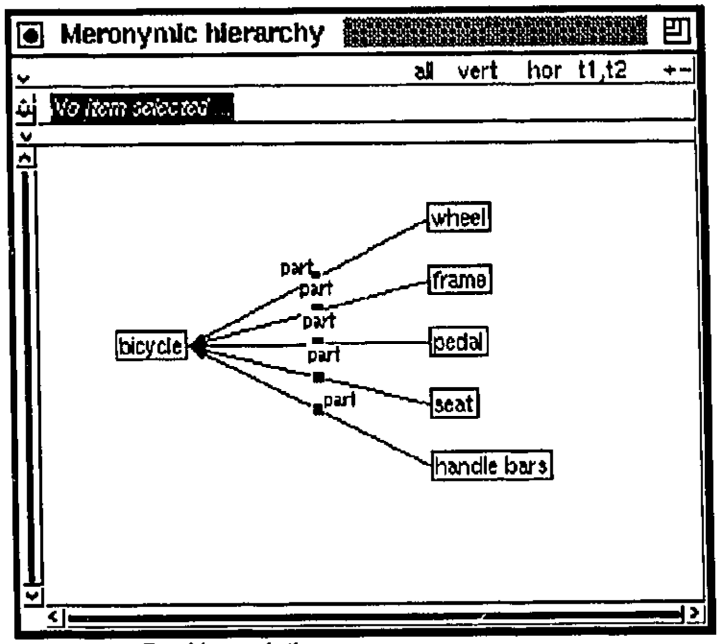


Figure 3. Partitive relations.

2. Associative relations

Figure 1 indicated that associative relations are non-hierarchical relations, and thus do not allow characteristics to inherit from one concept to the other. To the best of my knowledge, associative relations have not been studied in any kind of depth in the terminology literature; thus, no kind of definitive classification exists. Within the literature, however, many different types of associative relation are discussed. Sager [1990:24-25], for example, provides a rather lengthy list, including:

1. cause—effect (e.g. explosion—fallout)
2. process—product (e.g. weaving—cloth)
3. process—method (e.g. storage—freeze-dry)
4. activity—place (e.g. coalmining—coalmine)
5. object—operation (e.g. drill bit—drilling)
6. object—form (e.g. book—paperback)

B. Attributes

For the purposes of this thesis, attributes are not further classified. As was noted in Figure 1, they refer to the concept itself without involving any other concept.

Distinction 2: more important characteristics vs. less important characteristics

According to this distinction, characteristics have varying degrees of importance for definition construction. Once the information about a concept system has been recorded in the TKB, the terminologist can classify the characteristics according to their degree of importance for describing concepts.

The different types of characteristic discussed in section 1.2.3 imply that there is a "continuum of importance" for characteristics, with NECESSARY CHARACTERISTICS at one end and TYPICAL CHARACTERISTICS at the other. Necessary and typical characteristics play a major role in discussions about definitions. For this reason, they are examined in greater detail in chapter 4.

1.2.4 Inheritance of Characteristics

Inheritance is a powerful artificial intelligence technique that operates when concepts are arranged in a generic–specific hierarchy. It allows for any characteristic of a given concept to be implicitly true for all subconcepts of this concept, for all subconcepts of these subconcepts, and so on.

As concepts become more specialized, the value of a characteristic may either remain the same or undergo modification or specification (although the name will usually remain constant[16]). New characteristics may also be added at each new concept level.

Table 2 shows that the concept **CD–ROM** inherits certain characteristics from its

[16] Sometimes the characteristic name will change as it inherits from one concept to another. Miller [1993:116] provides an example of this phenomenon.

superconcept, **compact disc**, which in turn inherits characteristics from its superconcept, **optical disc**. Some of the characteristic values have remained the same, while others have been modified. For example, the characteristic *physical form:disc* does not change as it inherits from one level to another in the hierarchy. However, **optical disc**'s characteristic *encoding method:digital or analog* has been modified to *encoding method:digital* for **compact disc**. It can also be seen that an entirely new characteristic, that of *standard:one of Red Book, Yellow Book, Green Book, Orange Book*, is added at the level of **compact disc**.

characteristic name	*characteristic value*		
	optical disc	**compact disc**	**CD-ROM**
physical form:	*disc*	*disc*	*disc*
diameter:	*3.5, 4.72, 5.25, 8, 12, or 14 inches*	*5.25 inches*	*5.25 inches*
degree of writability:	*one of read-only, write-once, erasable*	*one of read-only, write-once, erasable*	*read-only*
encoding method:	*digital or analog*	*digital*	*digital*
storage capacity:	*varies with the type of optical disc*	*varies with the type of compact disc*	*up to 600 MB*
content:	*one or both of audio, data*	*one or both of audio, data*	*audio, data*
standard:	*n/a*	*one of Red Book, Yellow Book, Green Book, Orange Book*	*Yellow Book*

Table 2. Characteristics inheriting in a generic-specific hierarchy.

1.3 CONCEPT SYSTEMS

Obviously, concepts participating in relations do not exist in isolation: they combine with other concepts to form concept systems. A concept system is defined by ISO [1990a:5] as a "structured set of concepts established according to the relations between them, each concept being determined by its position in this set."

1.3.1 Representing Concept Systems

The larger and more complex a concept system is, the more useful it is to graphically represent this system in order to clarify the concept relations [ISO 1991:2]. Some graphical representations of concept systems include field diagrams (Figure 4), tree diagrams (Figure 5), and hierarchical lists (Figure 6).

				models
0 hand saw	0.1 non pre-stressed hand saw	0.1.1 carpenter's saw	0.1.1.1 hand saw for round timber	[picture]
		t.o.c.: cutting against the grain	0.1.1.2 hand saw for cut timber	[picture]
	t.o.c.:[17] to be held at one or both ends	0.1.2 cross-cut saw	0.1.2.1 long saw for timber	[picture]
		t.o.c.: blade-teeth laminated or not	0.1.2.2 link-tooth saw	[picture]
	0.2 pre-stressed hand saw	0.2.1. bow saw	0.2.1.1 tree saw	[picture]
		t.o.c.: cutting against the grains	0.2.1.2 bow saw for cut timber	[picture]
	t.o.c.: to be held at one or both ends	0.2.2 trimsaw	0.2.2.1 log saw for round timber	[picture]
		t.o.c.: cutting against the grains	0.2.2.2 log saw for square timber	[picture]
			0.2.2.3 pit saw	[picture]

Figure 4. Field diagram (Source: ISO 1988:12).

[17] t.o.c. = type of characteristics

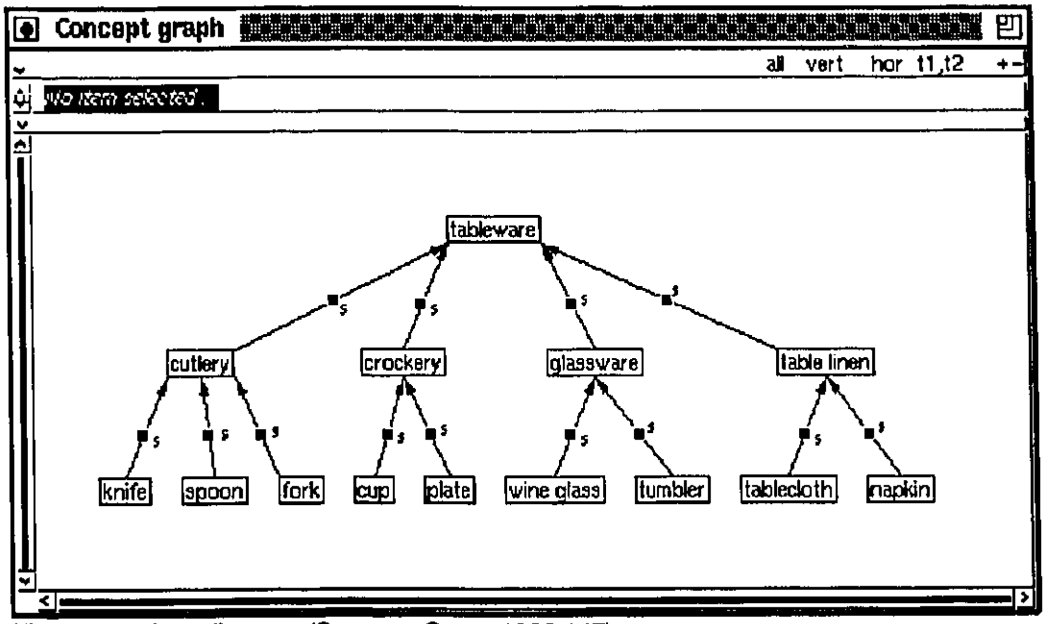


Figure 5. Tree diagram (Source: Cruse 1986:147).

As these figures demonstrate, graphical representations are particularly well-suited for depicting and identifying hierarchical relations. Unfortunately, as Skuce/Meyer [1990a:188–189] note:

living being	1
plant	1.1
cryptogamous plant	1.1.1
phanerogamous plant	1.1.2
gymnosperm	1.1.2.1
angiosperm	1.1.2.2
animal	1.2
protozoan	1.2.1
metazoan	1.2.2

Figure 6. Hierarchical list (Source: ISO 1991:5).

> ...it is still extremely rare to find in terminology banks or publications any systematic attempt to represent the conceptual organization of a field, whether textually[18] or graphically. Rather, one normally detects only traces of conceptual structures in the definition of certain terms, "somewhat like a puzzle that no one can put together because there are pieces missing, and there is no picture of the whole that can serve as a guide."[19]

[18] By "textually," Skuce/Meyer mean in the form of what ISO [1991] terms a hierarchical list [Ingrid Meyer, personal communication].

[19] Translated from Kulkuska-Holmes/Knowles [1989:382].

1.4 PRACTICAL USES FOR CONCEPT ANALYSIS

Concept analysis entails describing the attributes of individual concepts and the relations that hold within systems of concepts. Ideally, concept analysis within a given subject field results in a collection of characteristics for each concept, fully describing both the concept and its relations to other concepts. Many of these concept relations are best represented in graphical form, to help the user better understand how the concept system "hangs together."

Good concept analysis helps the terminologist carry out the practical operations of terminological work, such as

— ***delimiting and partitioning subject fields:*** by clearly understanding the concepts and concept relations of a given subject field, the terminologist can more easily delimit the subject field in which these concepts are found; conceptual "noise" (where concepts are mistakenly included in a subject field) and conceptual "silences" (where concepts appear to be missing from the subject field[20]) are more easily seen and remedied

— ***distinguishing quasi-synonyms:*** by comparing a list of characteristics for the concepts denoted by quasi-synonyms, the terminologist can verify whether or not the suggested synonyms actually denote the same concept

— ***neology:*** by looking at the resulting concept system, the terminologist can determine which concepts need terms created for them; she can also tell if a particular neology is unnecessary by comparing existing concepts with the one denoted by the new term

— ***communicating with subject-field experts and colleagues:*** by using the explicit information about a subject field's concepts and concept system, the terminologist can efficiently and

[20] For example, within the Typesetting TKB constructed during the pilot project at the Department of the Secretary of State of Canada, the terminologists included a relatively new kind of typesetting called **digital typesetting**. What appeared to be missing, however, was a kind of "non-digital" typesetting. By clearly understanding the subject field, the terminologists were able to see this silence and be on the lookout for a corresponding term (in this case, **analog typesetting**), which otherwise may have gone unnoticed [Ingrid Meyer, personal communication].

effectively discuss the subject field with experts and other terminologists

— ***interlingual equivalence (useful for translators):*** by comparing concepts denoted by terms in different languages via their characteristics, the terminologist is able to verify equivalence between these terms. As Cole [1987:78] notes, "it is only by ascertaining the correspondence of concepts employed in different languages that the terminologist can establish the equivalence between the terms used to designate them."

— ***constructing definitions (intensional definitions):*** by knowing a given concept's generic concept and by having a list of characteristics from which to choose those characteristics necessary to sufficiently describe the concept, the terminologist can feel confident that she is working with all the information needed to construct (or adapt or even simply criticize) a definition for that concept and its co-concepts.

Concept analysis plays a role in all aspects of terminology work, from subject-field delimitation to definition construction. While terminologists have always carried out concept analysis, they have been obliged to do so for the most part using what Meyer/Paradis [1991:3] refer to as "a combination of 'paper-and-pencil' and 'do-it-in-my-head' methods." It is only within the past few years that computerized tools have become available to facilitate this activity. One such tool, called CODE, is described in chapter 2.

Chapter 2 THE TECHNOLOGY: CODE

2.0 INTRODUCTION

CODE (Conceptually Oriented Description Environment) is a computerized, generic knowledge management system, developed at the Artificial Intelligence Laboratory of the University of Ottawa under the direction of Dr. Douglas Skuce.

The most recent version of CODE (CODE4), has been in use since mid-1992 (although it should be noted that certain features of CODE4 are still in the final stages of development). The previous version, CODE2, was widely used for both research and commercial purposes.[21] This thesis concentrates on CODE4, in contrast to two previous Master's theses involving the use of CODE, both of which used version 2 [cf. Bowker 1992, Miller 1993].

As a generic system, CODE is intended for a variety of applications, such as natural language processing, software specification and design, expert systems, teaching specialized subjects such as biology, and general terminological analysis [Skuce 1992:1]. To date, CODE2 has been used to develop terminological knowledge bases in the fields of typesetting [cf. Meyer/Paradis 1991, Skuce/Meyer 1990a/b], optical storage technology [cf. Meyer/Bowker/Eck 1992a/b, Meyer/Skuce/Bowker/Eck 1992a/b], and finance [cf. Miller 1993]. The optical storage technology TKB was recently converted to operate under CODE4.[22]

Because the focus of this thesis is a knowledge-based approach to terminology, CODE will be examined specifically from the viewpoint of how it can be used in a terminological application, i.e. to help the terminologist carry out concept analysis and build a TKB.

---

[21] CODE3 was an experimental prototype and, as such, not used for any actual applications.

[22] For the sake of simplicity, CODE4 is henceforth referred to as CODE.

2.1 CODE'S USER INTERFACE

CODE's user interface centres around the use of browsers, which are in effect windows which displaying the data (in the form of concepts, characteristics, and facets[23]) stored in a TKB. Browsers give the user a great deal of power and flexibility for viewing the TKB data, and are discussed in more detail in the next section.

CODE's interface is extremely versatile and user-friendly: a hypertext-like facility, using associative links, allows the user to navigate quickly and easily through the data. For example, someone who wants to quickly learn about concept X can select[24] it from a list of concepts in the concepts browser; X's characteristics then appear in the characteristics window. To view the characteristics of concept Y, the user selects concept Y in the concepts browser, and Y's characteristics appear in the characteristics browser. If the user then wants to see if concept Y shares characteristic A with concept X, she can first select characteristic A in the characteristic browser (concept Y is still selected) and then select concept X. If concept X also has characteristic A, A will continue appear in the characteristics browser (although it may have a different value).

Essentially, CODE's interface allows users to explore the TKB without having to know exactly what they are looking for; there is no central menu which must be referred to before data can be accessed. For those looking for a specific item of information, CODE's "goto" (search) feature can further aid their exploration.

2.2 CODE FEATURES

CODE possesses several features which can help terminologists carry out concept analysis,

[23] Facets provide information about a given characteristic (name and value). For example, the COGNITERM TKB has the facets (a) reference (to document the source of the characteristic value), (b) comment (to provide a general comment about the characteristic, if necessary), and (c) status (to show the degree to which the characteristic is "approved").

[24] It should be noted that CODE is mouse-driven. To *select* something means to click on it using the mouse.

specifically (a) a sophisticated textual and graphical display, (b) inheritance mechanisms, and (c) a flexible means of specifying characteristics. These features are described in detail below.

2.2.1 Textual and Graphical Display

CODE's browsers allow the user to view information in two main ways: textually or graphically. Information in a textual browser is seen in list form, while information in a graphical browser is seen in tree structure form. Browsers can be moved around the screen, resized, or even collapsed[25] as needed.

2.2.1.1 Textual Browser

The lists of data in textual browsers may be viewed either hierarchically (with hierarchy levels indicated by progressive indenting) or alphabetically. The user may restrict the list either by masking out (hiding) subhierarchies found within the main list, or by indicating that only a given subhierarchy should appear within the list. Figure 7 shows a textual browser containing a hierarchical list of concepts, with two collapsed subtrees (indicated by –+).

Browsers may work independently, or they may be created from another browser. In the latter case, the originating browser is said to be the driving browser and the resulting one the driven browser. A driven browser can either be attached to its driving browser or detached from it (i.e. appearing as a separate browser). If two or more browsers are attached to each other, they are considered sub–browsers within one window.

For example, the textual browser shown in Figure 7 can act as a driving browser in the creation of another textual browser containing conceptual characteristics (the driven browser). If a

[25] Depending on the environment that CODE is running in, browsers may either be collapsed into icons (similar to those of the Macintosh interface), or into a list item (as seen with the X–Windows' TWM Window Manager).

concept is selected in the driving browser, its characteristics are then displayed in the driven browser.

Figure 8 shows the concepts browser from Figure 7 (left) with an attached driven browser containing conceptual characteristics (right). Note that the characteristic value (*800–1000 MB*) of the selected characteristic (***storage capacity***) is displayed in the text pane (the small subwindow at the bottom right). In turn, the characteristics browser in Figure 8 may be used to create another browser containing facets; in this case, the characteristics browser is the driving browser and the facets browser is the driven browser. When a characteristic is selected in the characteristics browser, its facets appear in the driven facets browser. Figure 9 shows the window from Figure 8 with an additional attached facets browser. The

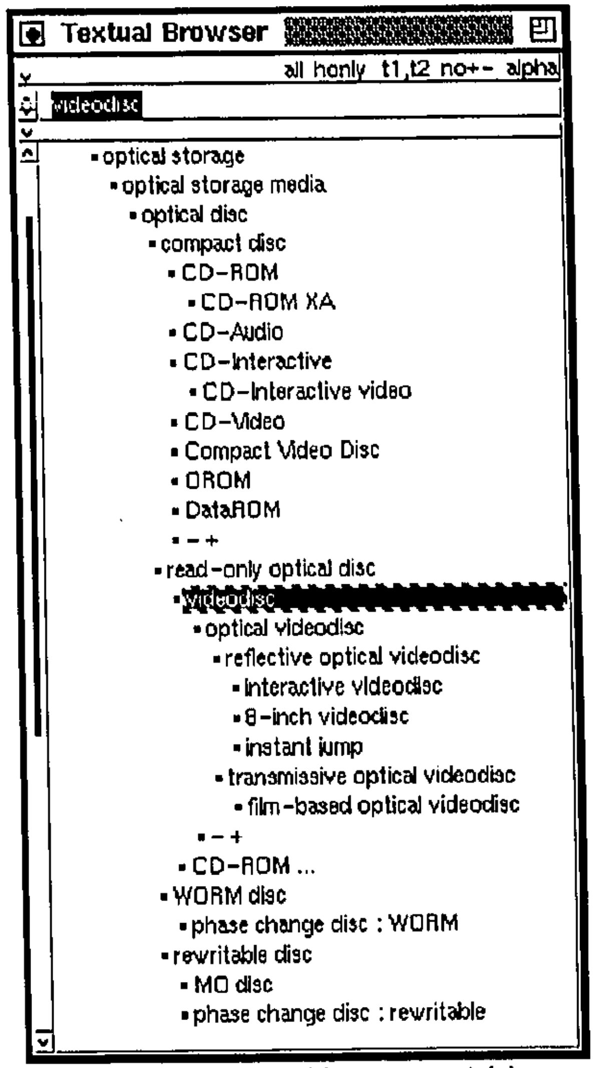


Figure 7. CODE textual browser containing a hierarchical list of concepts with two collapsed subtrees.

REFERENCE facet for ***storage capacity:800–1000 MB*** for the concept of **videodisc** has been selected; the facet value, *CHEN89B:4*, can be seen in the text pane.

It is possible that facets may be used more in the future to annotate each characteristic name and value. For example, the terminologist could record the source of characteristic name and the context in which it was found (e.g. a table of characteristics or a defining context). Chapter 6 discusses how this and other kinds of information may prove useful to the terminologist at a later date, for example, to explain the selection of certain characteristics to a subject-field expert. Another facet could then be used to record the expert's approval.

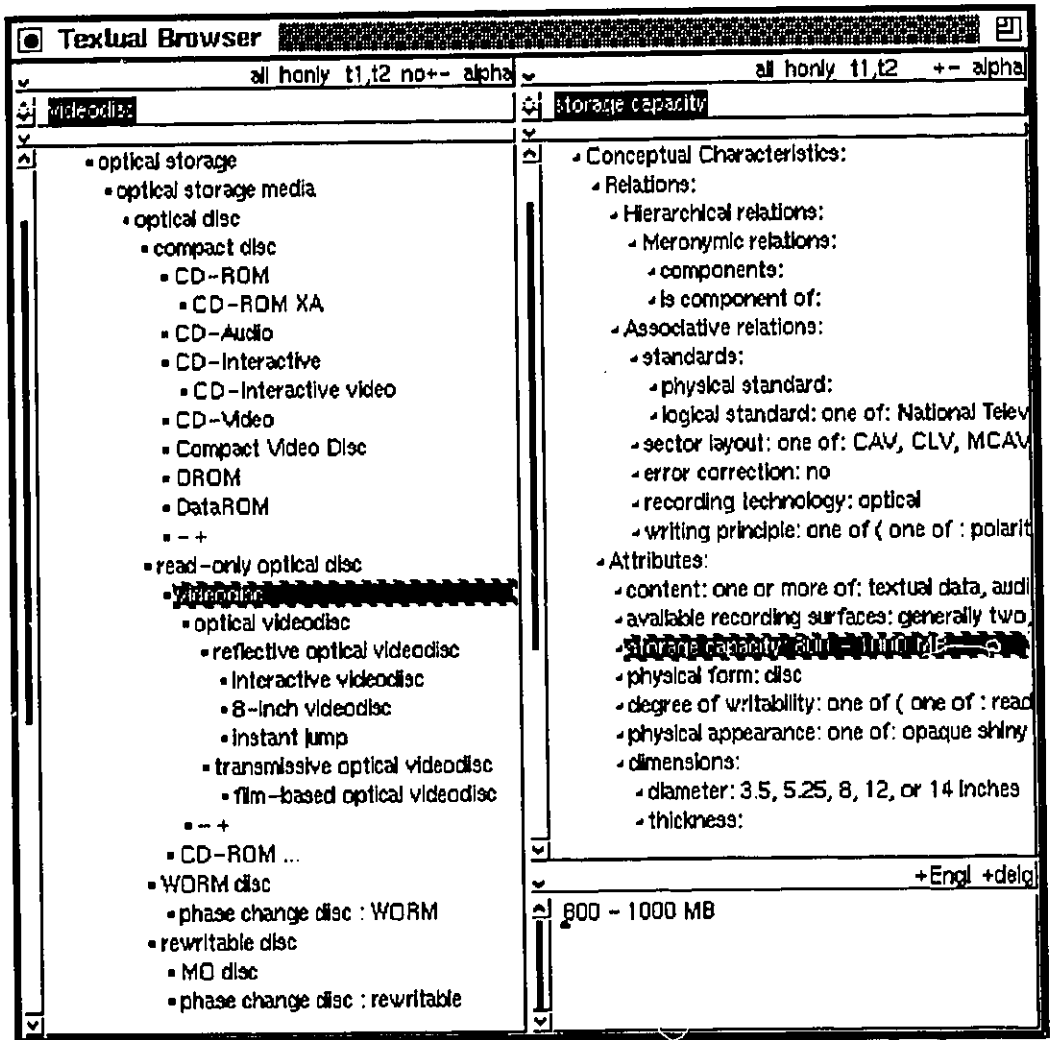


Figure 8. CODE window containing a concepts browser (left) and an attached browser containing a hierarchical list of characteristics (right).

Figure 7, Figure 8, and Figure 9 show how concepts, characteristics, and facets can all be easily viewed in textual browsers.

How the textual browser facilitates concept analysis

Textual browsers allow the terminologist to (a) methodically gather information about concepts

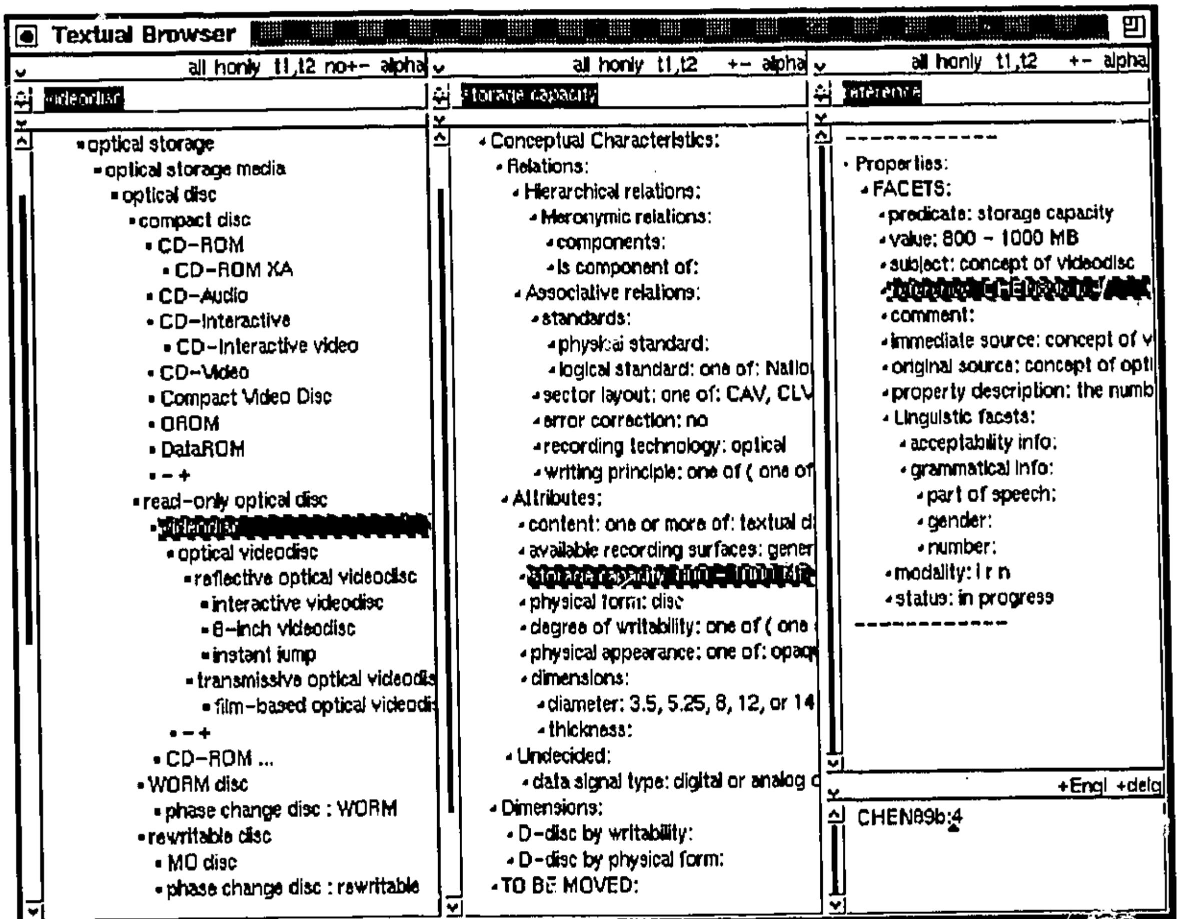


Figure 9. CODE window containing a concepts browser (left), an attached characteristics browser (middle) and an attached facets browser (right).

and (b) maintain control over this information. Lists of characteristics are available for each concept; these lists explicitly display all the information that has been entered for a given concept (or inherited from a generic concept). As noted above in section 2.2.1.1, textual lists may be alphabetical or hierarchical. Hierarchical lists of concepts show relations[26] between concepts to a certain degree (the graphical display, discussed next, is more effective at showing these relations). Characteristics may also be grouped into hierarchies, if desired, to display them more logically. Alphabetized lists of both concepts and characteristics help the terminologist quickly locate data.

[26] It should be noted that, although generic–specific relations are automatically shown in CODE, other types of relation may also be displayed.

2.2.1.2 Graphical Browser

The graphical browser (commonly referred to as the graph) provides a well-developed graphical representation of all the concepts in the TKB and the relations between them. Although the graph automatically shows generic-specific relations between concepts, partitive and associative relations may also be displayed.

Figure 10 shows a simple graphical browser with generic-specific relations showing. On the graph, *links* (arrowed lines representing the concept relation) are used to connect *nodes* (representing concepts). These links are labelled "s" by default, indicating the generic-specific relationship between concepts. Other hierarchical and associative relations may also be represented on the graph. Figure 11 shows a graph with generic-specific, partitive (part) and associative (here: device) links. Although typically it is concepts which are displayed graphically, characteristics may also be viewed in this manner.

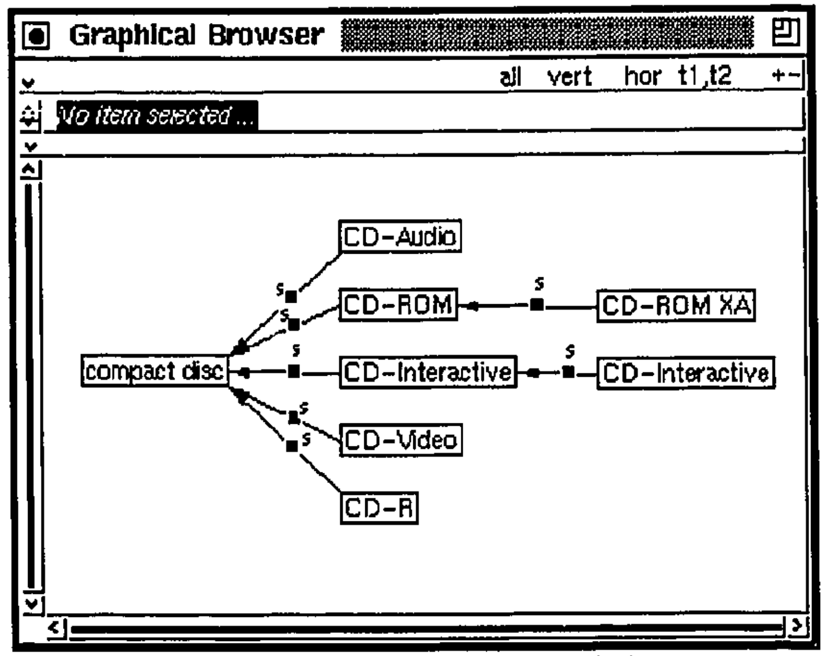


Figure 10. CODE graphical browser from the optical storage technology TKB displaying generic-specific relations.

As with textual browsers, a graph may display the whole TKB hierarchy or it may be restricted to certain concepts. Subtrees may be collapsed or a particular subtree may be shown, allowing the user to concentrate on those particular concepts. Figure 12 shows a subtree of the optical storage technology TKB with a collapsed subtree and node.

One type of graph in CODE, called the characteristic link graph, automatically shows a

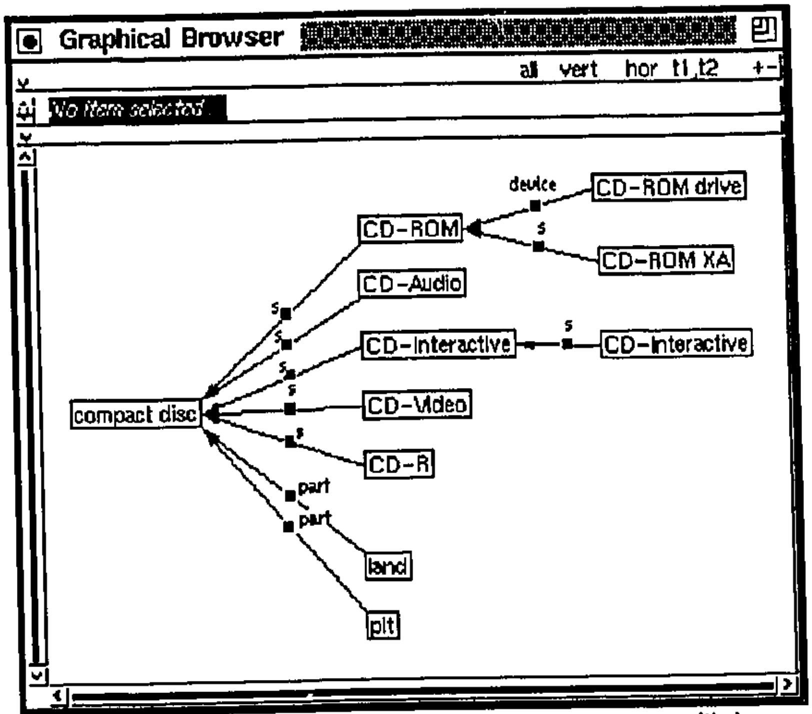


Figure 11. CODE graphical browser showing concepts with isa (s), partitive (part) and associative (device) links.

concept's associative relations, showing how a given concept is linked by non-isa relations to other concepts in the TKB. Figure 13 shows a characteristic link graph for the concept **CD-ROM**, from the optical storage technology TKB.

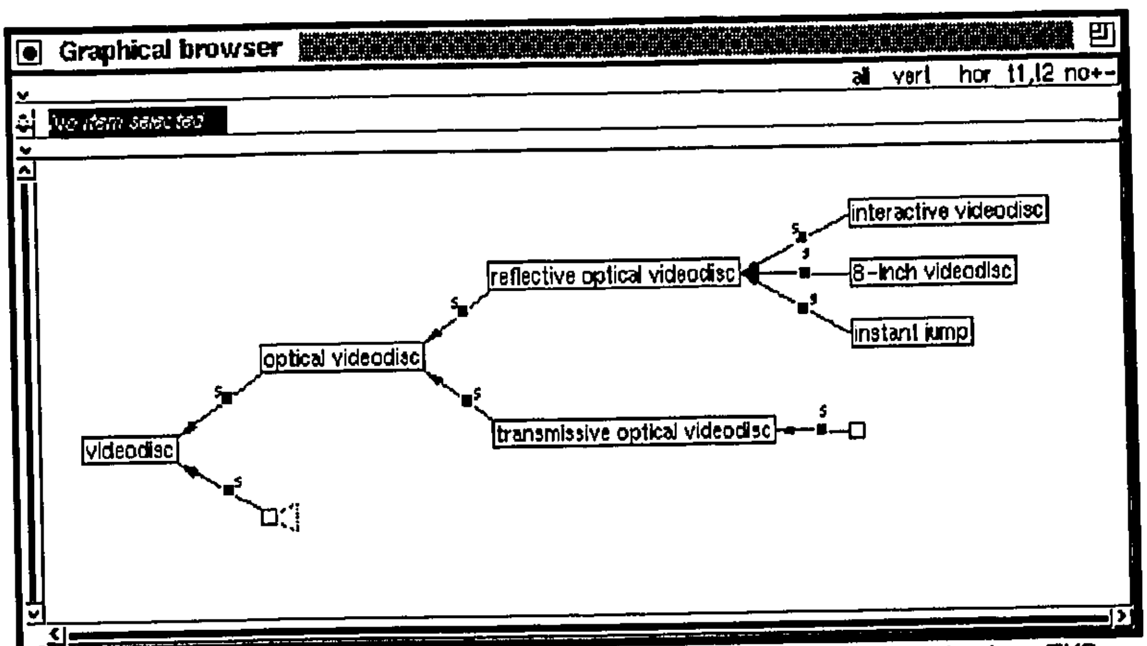


Figure 12. CODE graphical browser showing subtree of the optical storage technology TKB, with one collapsed node (indicated by small square) and one collapsed subtree (indicated by fan shape).

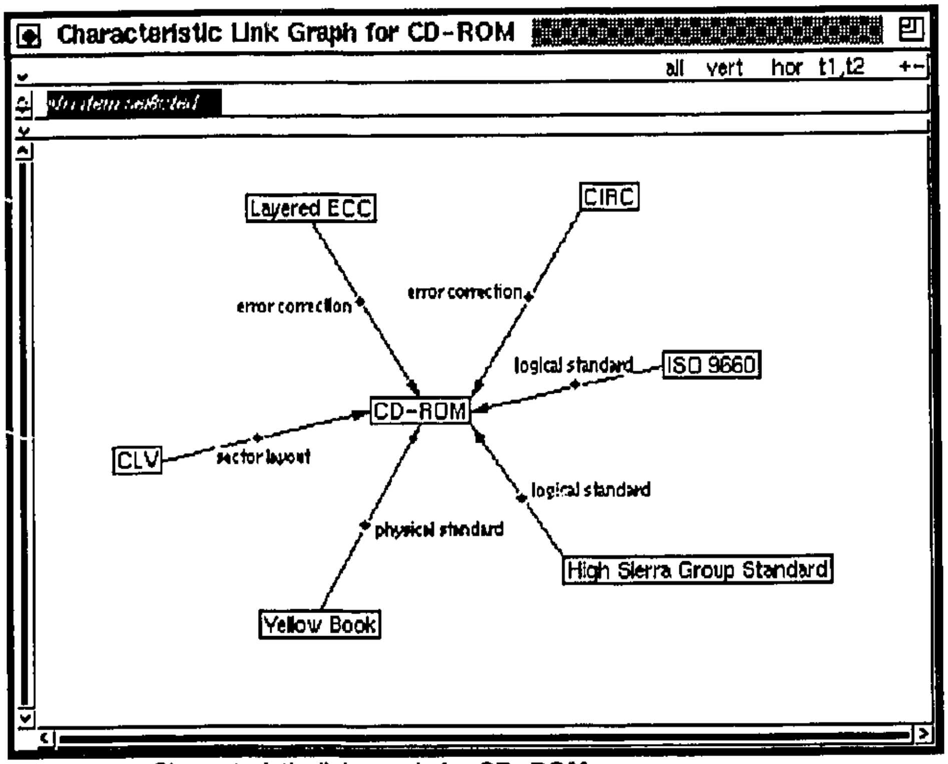


Figure 13. Characteristic link graph for **CD-ROM.**

The concepts linked to **CD-ROM** denote values for some of the associative relations (characterisdcs) of **CD-ROM**; each link label denotes an associative relation (characteristic) name.

How the graphical browser facilitates concept analysis

Visual representation is often the most effective and rapid method of conveying information. For example, the graphical browser is especially useful in providing users with an overview of the entire subject field. Even using the pencil-and-paper method, terminologists often find it helpful to draw concept networks of their selected fields or subfields. As Skuce/Meyer [1990a:4] note:

> We cannot overemphasize the importance of good graphical representations for terminology: in our experience,[27] the terminologists referred to the graph constantly and wanted the graph window open at all times, no matter what other components of the knowledge base were being focused on.

[27] Using CODE during a pilot project at the Department of the Secretary of State of Canada.

The graphs showing associative relations make it easy for the terminologist to see how different areas of the subject field are related, and to make decisions about what kind of information should be found in the characteristics browser for each concept.

2.2.1.3 Combining Browsers

The user is not restricted to having one kind of browser per window. One window may in fact contain any combination of graphical and textual browsers. Figure 14 shows a window in which concepts are displayed in graphically while conceptual characteristics are displayed textually.

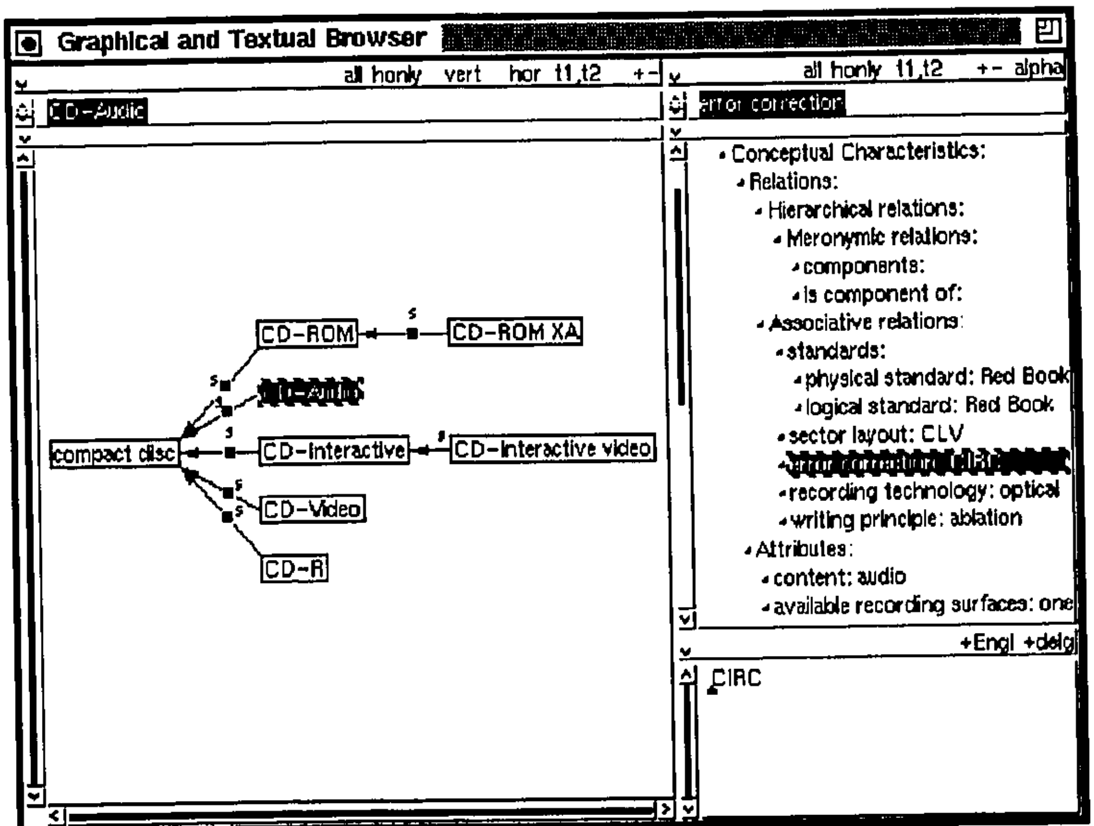


Figure 14. CODE window with graphical concepts sub-browser (left), and textual characteristics sub-browser (right).

2.2.2 Inheritance Mechanisms

As was explained in chapter 1 (section 1.2.4), inheritance is a powerful AI technique that can operate when concepts are arranged in a generic–specific hierarchy. It allows any characteristic of a given concept to be implicitly true for all subconcepts of this concept, for all subconcepts of these subconcepts, and so on. CODE allows the user to arrange concepts in inheritance hierarchies, so that more specific concepts may inherit characteristics from more general ones.

How inheritance mechanisms facilitate concept analysis

When terminologists identify a given concept as a member of a known set of concepts, they can automatically access this additional (inherited) knowledge about the concept. This ensures that the information recorded for related concepts is consistent, and that all subconcepts of a concept have similar characteristics.

CODE's inheritance mechanisms also help the terminologist reduce the amount of work needed to input the same characteristics for many related concepts. Without these inheritance mechanisms, the terminologist would be compelled to input the characteristic names and values for each and every concept, a time–consuming and repetitive task, and one that increases the likelihood of an input error.

CODE's inheritance feature is complemented by other mechanisms to ensure that if the terminologist attempts to include information which should not logically be there, the error is pointed out (in the form of a dialog box). For example, referring to Table 2 (chapter 1, section 1.2.4), which shows inherited characteristics of concepts, if the user were to change the characteristic ***encoding method:digital*** *or analog* of **optical disc**, she would be signalled that this would create an inconsistency with **compact disc** and **CD–ROM** (both of whose characteristics have been modified to ***encoding method:digital***). The user could then choose to have the new characteristic value inherit to **compact disc** and **CD–ROM** (overwriting the current value) or perhaps rethink the intended value change at **optical disc**.

2.2.3 Flexible Means of Specifying Characteristics

Since CODE does not presuppose any particular theory of concepts, the terminologist can create any number of characteristics and any type of characteristic, arranged hierarchically (as in Figure 15) or not. Concepts may have characteristics (conceptual characteristics), as may terms denoting concepts (linguistic characteristics). As noted in section 2.2, even characteristics may have characteristics (facets).

The only characteristic which CODE provides by default is the generic–specific (or isa) relation, which it considers to be the primary relation between two concepts. If the main relation between two concepts is *not* generic–specific, then CODE provides other ways of linking these concepts according to their main relation.

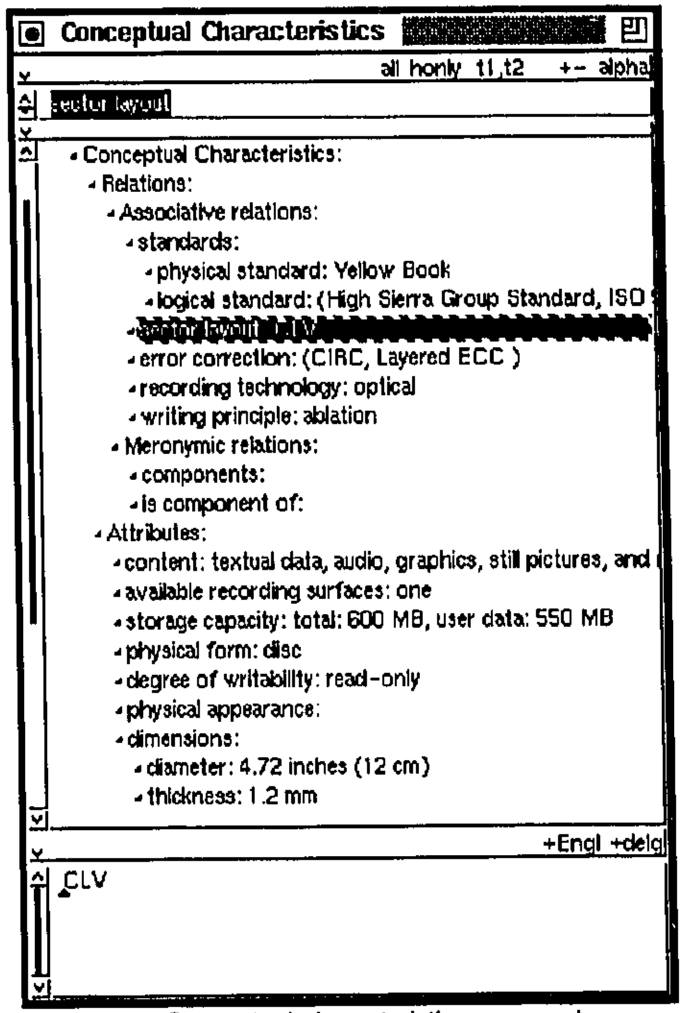


Figure 15. Conceptual characteristics arranged hierarchically for the concept **CD–ROM.**

How flexible means of specifying characteristics facilitates concept analysis

CODE's flexibility ensures that the terminologist can tailor characteristics in the TKB to best suit the subject field under investigation. Any number of characteristics and any type of characteristic may be used to present the information about a given concept, its term, and their respective characteristics. As mentioned earlier, these characteristics may be placed in hierarchies which display the information more clearly, or they may be left in non-hierarchical lists.

Concept relations that are not generic-specific may be specified by the terminologist as conceptual characteristics. These relations (whether partitive or associative) may then be displayed on the graph along with the regular generic-specific relations; on a graph which shows only one particular relation, such as a partitive hierarchy; or viewed one concept at a time in a characteristic links graph (Figure 13).

As a generic knowledge management tool, CODE provides terminologists with a systematic environment for carrying out concept analysis using a knowledge-based approach. It helps terminologists gather, organize, and retrieve the knowledge they acquire about concepts and their characteristics using a variety of features, and to do so in a more systematic and explicit manner than is possible with pencil-and-paper or 'do-it-in-my-head' methods.

To effectively use a tool such as CODE, however, terminologists need guidelines. A general approach to concept analysis, developed by the COGNITERM team to create terminological knowledge bases, provides such guidelines. This approach is described in the next chapter.

Chapter 3 THE APPROACH: GENERAL CONCEPT ANALYSIS

3.0 INTRODUCTION

Chapter 2 explained how terminologists must follow an approach of some kind if they are to effectively use a knowledge management tool such as CODE to carry out concept analysis. This chapter discusses a general approach to concept analysis that helps terminologists use CODE within the framework of a terminology research project. It was developed by a team of COGNITERM researchers who, at the same time, created a TKB for the subject field of optical storage technology, concentrating on the subfield of optical storage media.

One key reason why this approach may be considered general is that it does not provide any specific guidelines for definition construction: it is restricted to the collection of information about concepts in a particular subject field (in the form of conceptual characteristics) and their respective terms (in the form of linguistic characteristics). It was designed generically, with the assumption that certain steps would have to be modified or added to ensure that definition construction could be effectively carried out once all the information about a concept had been gathered. Modifications and additions to this general approach, made with a view to definition-oriented concept analysis, are discussed in part 2 of this thesis.

To lay the groundwork for part 2, this general approach is discussed in detail below, using the extended example of **CD-ROM** to explain what steps the terminologist should take to use CODE to carry out general concept analysis and create a COGNITERM TKB.

First, however, I describe the general structure of COGNITERM, i.e. how the browsers are used to display the information gathered about a concept.

3.1 THE STRUCTURE OF COGNITERM

COGNITERM is conceived as a hybrid of a term bank and a knowledge base. It is most easily understood as having two main components: one linguistic (corresponding to the term bank) and one conceptual (corresponding to the knowledge base). There is also a third component, as seen in Figure 16, which provides valuable information about other aspects of the concept: background information. When a terminologist has fully researched a concept and entered these three kinds of information into COGNITERM, the result is a complete and thorough presentation of what the terminologist knows about the concept.

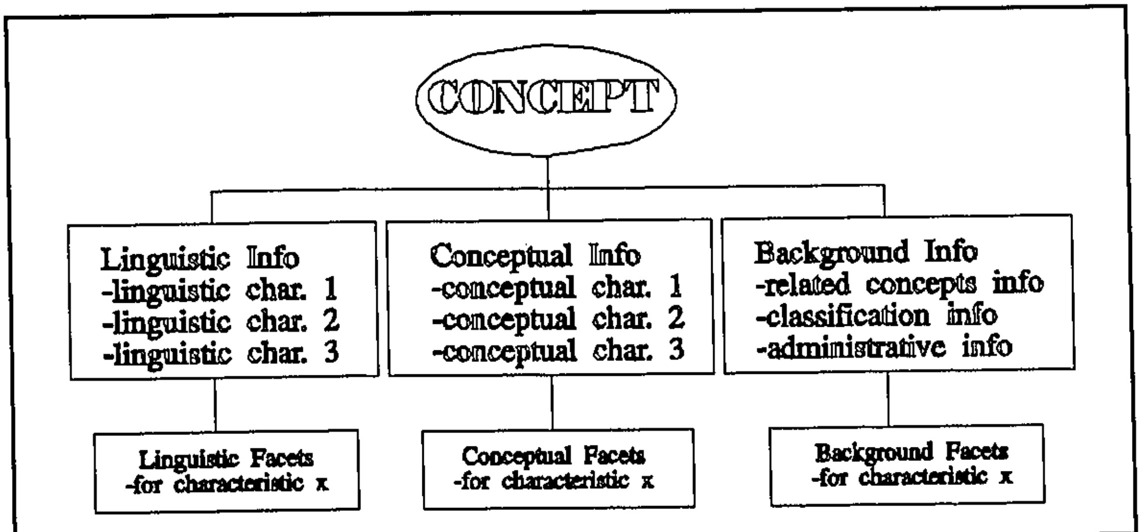


Figure 16. The Structure of COGNITERM.

In COGNITERM, each component is given its own separate "Information" browser, in which the information (i.e. Linguistic, Conceptual, or Background) is provided in the form of characteristics. In turn, each Information browser has an attached facets browser; facets, as mentioned in chapter 2 (section 2.1), provide further details about a given characteristic.

As was noted in chapter 2, CODE is very flexible in the way that it allows information to be displayed; browsers can be either textual or graphical. The following browsers were created for the COGNITERM TKB interface.

3.1.1 Concepts Browser

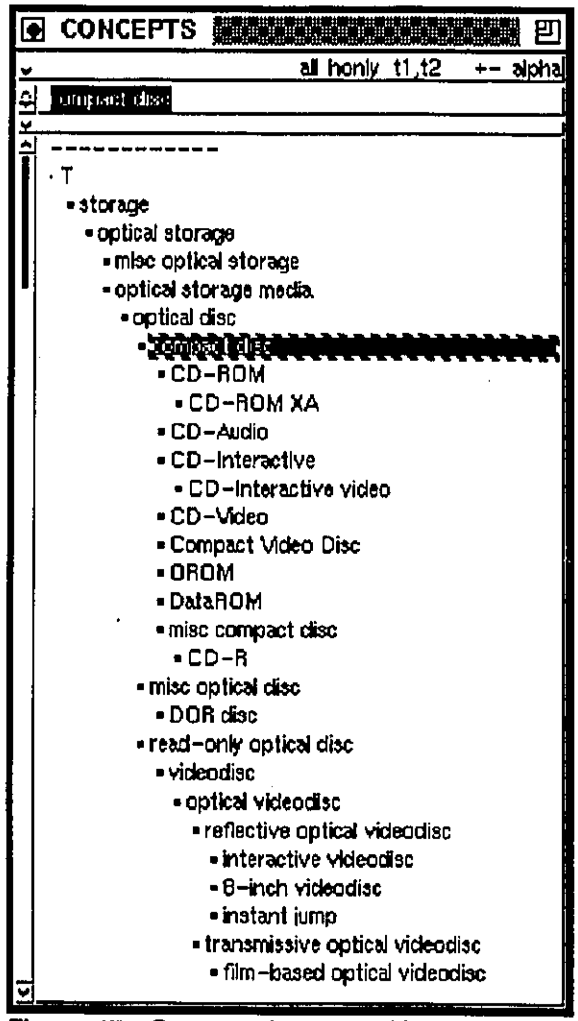


Figure 17. Concepts browser with concepts displayed hierarchically. The concept **compact disc** is selected.

The Concepts browser, displayed in Figure 17, provides the terminologist with a list of concepts in the TKB. Concepts can be viewed either alphabetically or hierarchically (as shown in Figure 17) by clicking on the *alpha/hiera*[28] button located at the top right corner of the browser. Individual subtrees can also be viewed within the browser.

From this browser, the terminologist can select any concept; the name of the selected concept is then displayed in the editing area (top left corner of the browser). In Figure 17, **compact disc** has been selected. As noted in section 3.1, three different types of information are available for the selected concept. This information is found in the Conceptual Info browser, the Linguistic Info browser, and the Background Info browser, each of which is discussed below.

3.1.2 Conceptual Info Browser

For COGNITERM's conceptual component, every concept is provided with a set of characteristics. Characteristics may be inherited (fully or partially) from more generic concepts or they may be added to the selected concept.

[28] Because the alpha/hiera button is a toggle button, only one selection appears onscreen at a time.

The Conceptual Info browser is driven off the Concepts browser. It is composed of two browsers: the conceptual characteristics browser and the conceptual facets browser.[29] Figure 18 shows the conceptual characteristics browser (left) displaying a list of conceptual characteristics (***Relations*** and ***Attributes***) describing the concept selected in the Concepts browser. The conceptual facets browser (right) in turn displays the facets (e.g. *VALUE* and *REFERENCE*) that describe the characteristic selected in the conceptual characteristics browser.

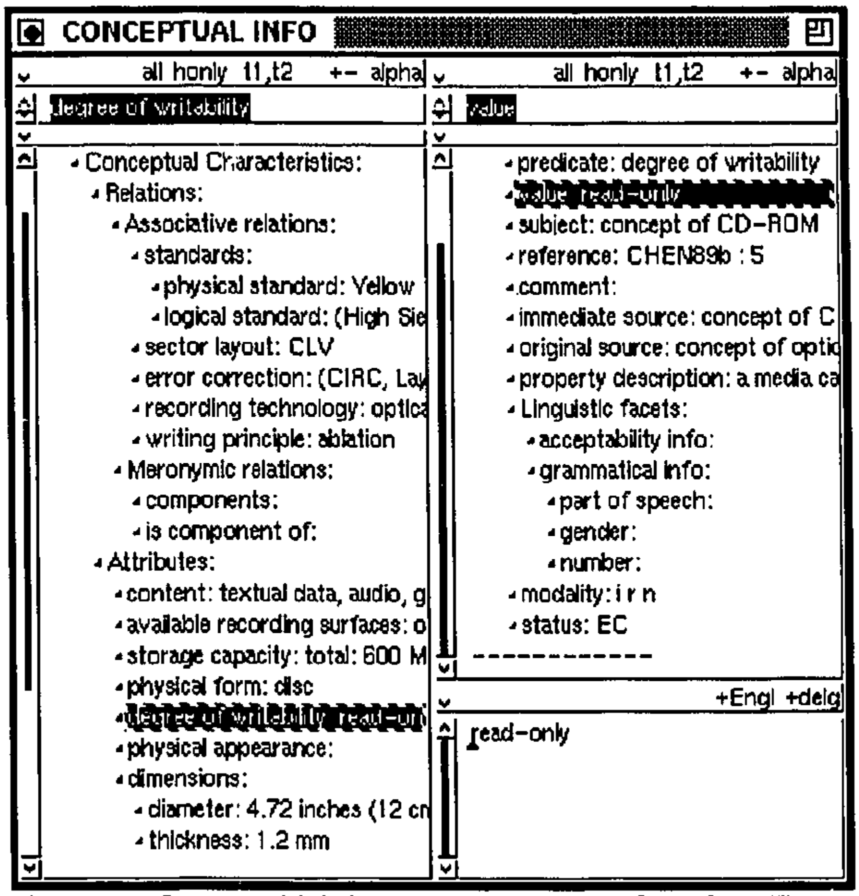


Figure 18. Conceptual Info browser for the concept **CD-ROM**. The characteristic ***degree of writability*** is selected (left), along with the value facet (right).

[29] All of the "Info" browsers (Conceptual, Linguistic, and Background) have this dual browser format, with characteristics in one browser and facets in another.

The text pane (bottom right) shows the value of the selected facet. For example, as seen in Figure 18, the selected conceptual characteristic (belonging to the concept **CD-ROM**) is ***degree of writability:read-only***. Since the *VALUE* facet is selected, it is the value of the characteristic that appears in the text pane. If the *REFERENCE* facet had been selected, the reference of the characteristic (here: *CHEN89B:5*) would appear in the text pane. Facets are described in further detail in section 3.1.5.

Characteristic hierarchies

With CODE4, it is possible to place conceptual characteristics within hierarchies. In COGNITERM, not only are there different kinds of characteristic (i.e. ***Relations*** and ***Attributes***), there are more specific kinds of ***Relations*** and ***Attributes***. As seen in Figure 18, ***Relations*** is further divided into ***Associative relations*** and ***Partitive relations***. For example, under ***Associative relations***, there are specific kinds of ***standards*** (***physical standard*** and ***logical standard***), while under ***Attributes***, there are specific kinds of ***dimensions*** (***diameter*** and ***thickness***). Several of the top levels of these characteristic hierarchies are essentially "nil" characteristics: as they serve to simply "describe" the characteristic hierarchy, they have no value. For example, the top-level Associative relation ***standards*** has no value, but its subcharacteristics, ***physical standard*** and ***logical standard***, do have values.

3.1.3 Linguistic Info Browser

For COGNITERM's linguistic component, the goal has been to include, for every concept, all the linguistic information that would normally be found in a conventional term bank,[30] plus some additional information such as collocations.

This information is displayed in the Linguistic Info browser. As Figure 19 shows, the

[30] The term bank that the COGNITERM Project used as a guide was the Government of Canada's TERMIUM III.

linguistic characteristics browser (left) displays characteristics which describe the term denoting the concept selected in the Concepts browser. The facets browser (right) displays facets describing the selected linguistic characteristic (e.g. *VALUE*, *REFERENCE*, *GRAMMATICAL INFO*, *ACCEPTABILITY INFO*). The characteristics are hierarchically grouped: the two main hierarchies are English Information and French Information[31] (although any language could be included), with each of them further divided into Term, Synonyms, Phraseological contexts, Defining contexts, and Derivates.

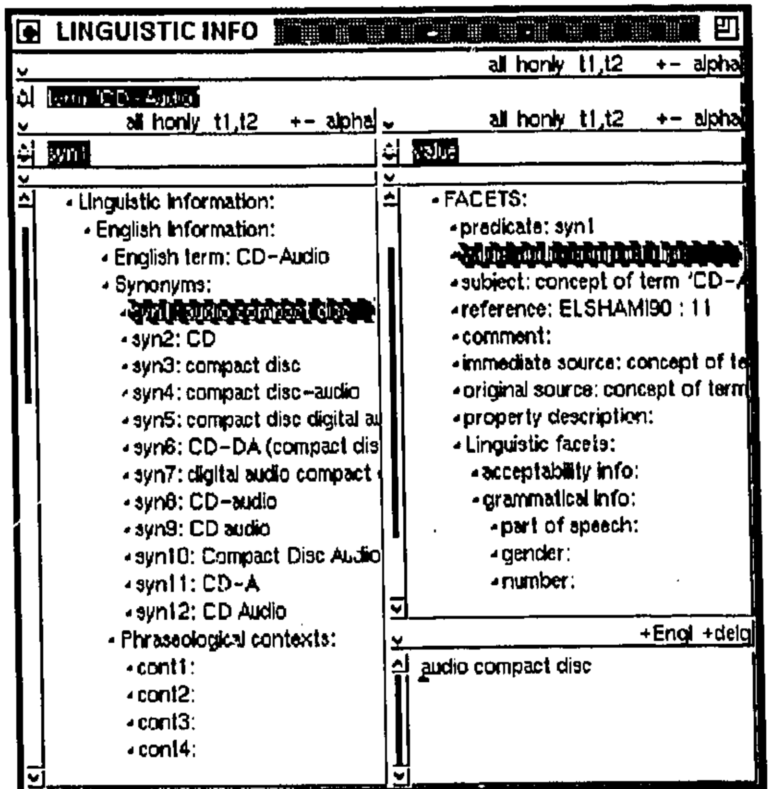


Figure 19. Linguistic Info browser for the concept **CD-Audio**. The characteristic ***syn1*** is selected (left), along with the value facet (right).

Some comments are found below about the four different types of linguistic information:

— ***Synonyms:*** COGNITERM provides for up to 12 synonyms[32] per language to be displayed in the Linguistic Info browser. The terminologist should be fully satisfied that the synonym is indeed an alternative term used to designate the concept and not a pseudo- or quasi-synonym. On a superficial level, this can be done through documentary evidence and expert opinion. On a more substantial

[31] Because of the number of characteristics in this browser, French Information is not displayed in Figure 19. CODE provides scrolling capabilities which allow the user to access information that is not immediately seen in the browser.

[32] This is the default provision; more can be added if the user desires. The same holds for phraseological contexts, definitions, and derivates.

level, the terminologist can compare the concepts designated by each term (through their characteristics) to see whether they are the same.

— ***Phraseological contexts:*** Phraseological contexts serve to provide collocational information. (It should be noted that collocations are not provided explicitly in TERMIUM III, although they are sometimes found in TERMIUM's textual support pane.) COGNITERM provides for 6 phraseological contexts per language.

— ***Defining contexts:*** Any defining contexts that the terminologist finds during her initial research are included in this subhierarchy. COGNITERM provides for 6 defining contexts per language. When COGNITERM was initially set up, it was decided that a place for defining contexts should be included because there was as yet no approach for constructing definitions. It was thought that these defining contexts might help later with definition construction.

— ***Derivates:*** Derivates serve to provide the terminologist with information about non–nominal terms that describe the concept (e.g. verbs, verb phrases, adjectives, adjective phrases). COGNITERM provides for 3 derivates per language.

3.1.4 Background Info Browser

The Background Info browser contains other information about a selected concept, as seen in Figure 20:

— ***Related concepts:*** the concept's generic concept(s) and specific concept(s), etc.

— ***Classification info:*** information about the concept's classification status. In COGNITERM, concepts may be classified or unclassified, depending on how much is known about them.

— ***Administrative info:*** "housekeeping" information such as the TKB name, the administrative status of the concept (e.g. *complete* or *in progress*), whether there is graphical support for the concept, the names of people who worked on the concept (***changers***), etc.

— ***Non-inheriting characteristics:*** conceptual characteristics which are pertinent only to that concept (they are placed here so as to not inherit to other concepts); for example, as seen in Figure 20, **CD-Audio** has the non-inheriting characteristic ***introduction date:1979***.

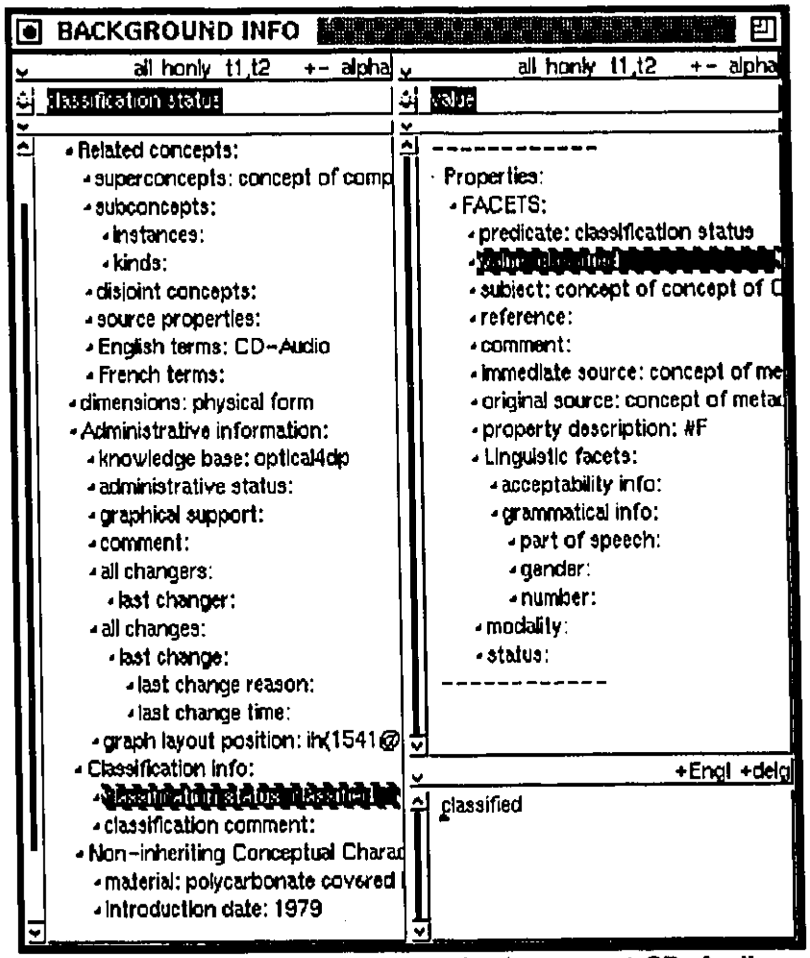


Figure 20. Background Info browser for the concept **CD-Audio.** The characteristic ***classification status*** is selected (left), along with the value facet (right).

3.1.5 Facets Browsers

Facets provide information about individual characteristics, as mentioned above. Each individual characteristic, whether linguistic, conceptual, or background, has a collection of information about it in the form of facets.[33] These facets are found in their own browsers attached to their respective Info browsers. Each kind of facets browser contains the same set of facets, although there are some additional facets for linguistic characteristics. The most important facets for terminologists (seen in Figure 21) are the following:

— ***NAME:*** the name of the characteristic

[33] Facets are in fact themselves characteristics (i.e. characteristics of characteristics), but for clarification they have been given another name.

— *VALUE:* the value of the characteristic

— *REFERENCE:* the source of the information about the characteristic value (all bibliographic information is included in a special "dummy" concept called **References**)

— *COMMENT:* any note that the terminologist wishes to include about the characteristic (typically, regarding its value)

— *CHARACTERISTIC DESCRIPTION:* a description of the characteristic name; e.g. for the conceptual characteristic ***error correction***, the description provided is *A SYSTEM OF ONE OR MORE CODES ON THE DISC, USED TO CORRECT MANUFACTURING ERRORS AND READING ERRORS RESULTING FROM SCRATCHES OR PARTICLES ON THE DISC.*

— *ORIGINAL SOURCE:* the concept where the characteristic was first introduced

— *IMMEDIATE SOURCE:* the concept where the characteristic value was last modified

— *LINGUISTIC FACETS* (for linguistic characteristics only; shown in Figure 22)

— *ACCEPTABILITY INFO:* the acceptability

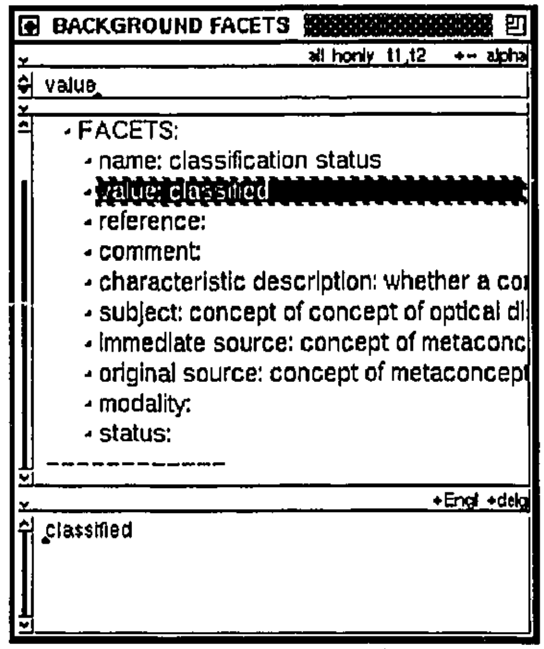


Figure 21. Background facets browser with *VALUE* facet for ***classification status:classified*** selected (for **optical disc**).

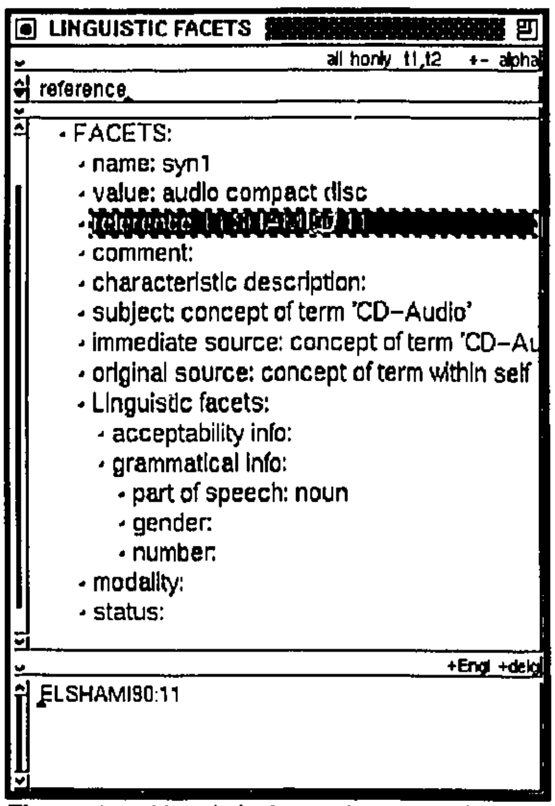


Figure 22. Linguistic facets browser with reference facet for ***syn1:audio compact disc*** selected (concept of **CD-Audio**).

code for the term (e.g. *APPROVED* or *DO NOT USE*)

— *GRAMMATICAL INFO:* grammatical information about the term, including *PART OF SPEECH*, *GENDER*, and *NUMBER*.

3.1.6 Full Graph

The Full Graph is a graphical browser which displays the same information as the Concepts browser. Generic–specific relations are shown by default, while other types of relation (e.g. partitive relations and associative relations) can be made to appear as well. Figure 23 provides an example of a full graph.

3.1.7 Characteristic Link Graph

The Characteristic Link Graph shows all non–generic–specific relations between a given concept and other concepts in the TKB. While it is not always in view, it can be easily made to appear. It is useful for ascertaining how a concept is related to other concepts within the subject field. Figure 13 in chapter 2 shows a characteristic link graph for the concept **CD–ROM**.

Together, these browsers allow the terminologist to (a) monitor the information about a concept as it is gathered and (b) explicitly show the final results of her concept analysis, whether through the concepts and their relations displayed in the Concepts browser and on the Full Graph, or through the characteristics shown in the various Information browsers.

The steps that the terminologist should follow to carry out concept analysis within the framework of a terminology research project are discussed in the next section.

3.2 STEPS IN CONCEPT ANALYSIS

While no two terminology research projects are the same, given the numerous and varied elements that are involved in each (e.g. subject field chosen, number of languages involved, number of researchers involved), some general key steps need to be followed in order to carry out concept analysis using CODE. These steps, which are standard practice in terminology, are described below, accompanied by the extended example of **CD-ROM** in the subject field of optical storage technology, and more specifically the subfield of optical storage media.

3.2.1 Selecting an Introductory Documentation Corpus

Once the subject field has been decided upon, as well as the number of languages that the project will involve, the terminologist must become acquainted with the subject field. To do this, she must compile and read an introductory corpus containing documentation that provides an overview of the subject field.

For example, once optical storage technology had been selected as the subject field, and English and French as the two languages,[34] the COGNITERM team chose a small corpus of documentation on optical storage technology, including monographs and periodical articles.

3.2.2 Establishing a Preliminary Concept Graph

After some introductory reading, the terminologist should have enough information to draw a preliminary concept graph using CODE. While information is most easily added to the TKB via the various textual browsers, it is the graph which best represents the knowledge thus far obtained. As Meyer [1991:14] notes, it has been shown that "terminologists place an extremely high value on graphical representations of knowledge structures." This initial concept graph delimits the various subfields that exist within the subject field.

[34] For simplicity, I concentrate on English only within this thesis.

Figure 23 shows part of the preliminary concept graph drawn for the optical storage TKB (this

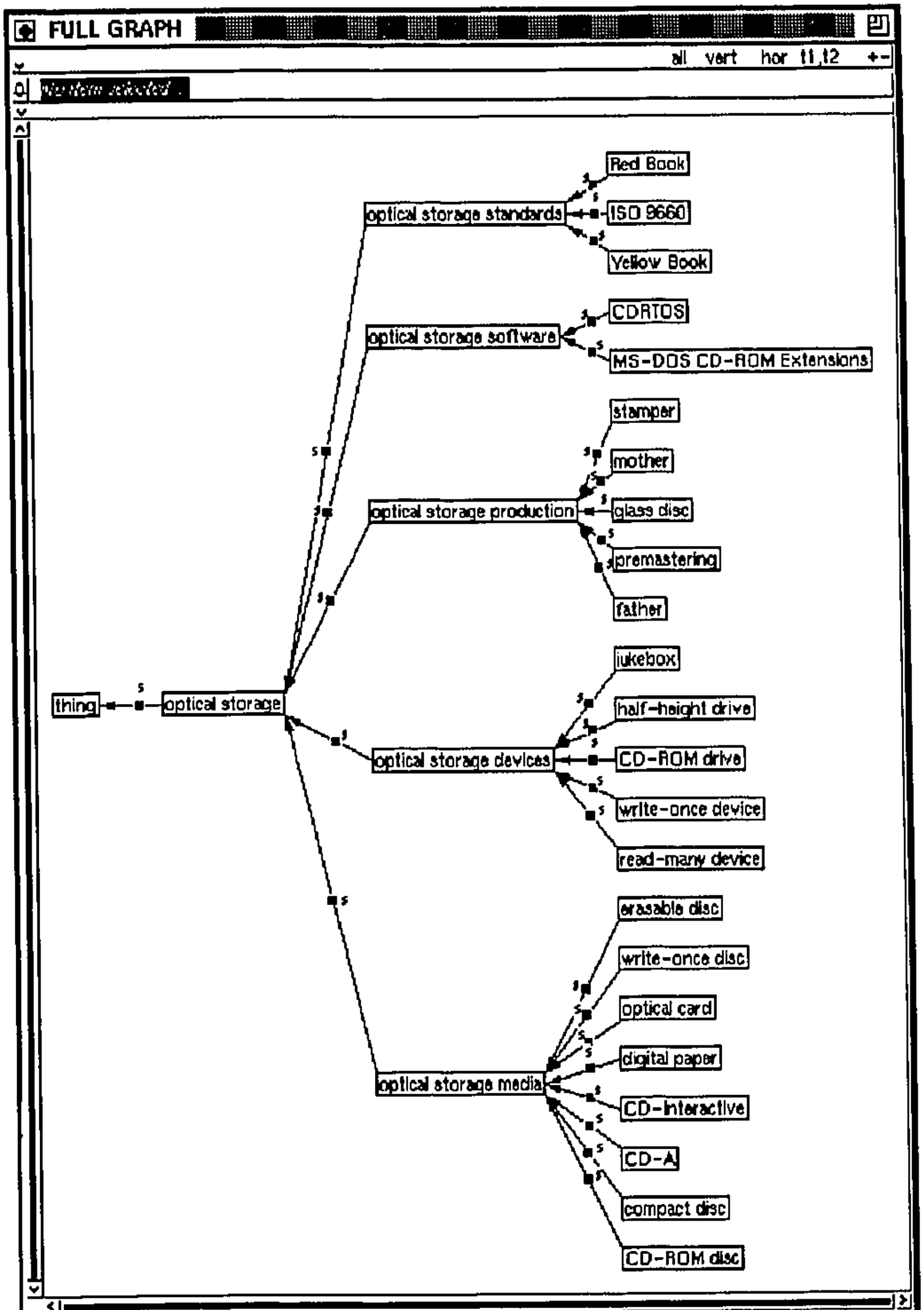


Figure 23. Preliminary concept graph for optical storage technology TKB with "concept clusters."

graph evolved into the full graph, showing the entire hierarchical structure of the subject field). Within the subject field of optical storage technology, it was seen that there were several subfields, in the form of "concept clusters"[35] which warranted further attention. These subfields were optical storage devices, software, standards, production, and media. We decided that we would initially explore the subfield of optical storage media, since it appeared to be not only fundamental to understanding the other subfields but, fortunately, the most tractable as well [Meyer/Bowker/Eck 1992b:4].

The graph also helps the terminologist to establish links with related documentation for a more extensive and focused documentation search, and to communicate with documentation resource persons. For example, we knew from our preliminary graph that we wanted to search principally for documentation on optical storage media, and were able to effectively explain to librarians and subject-field experts what we were searching for.

3.2.3 Compiling a Bibliography

Once the principal documentation has been assembled on the basis of the introductory documentation and suggestions from experts and documentation resource persons, the terminologist should compile a bibliography.

The bibliographical information is entered directly into COGNITERM, specifically to the **References** "dummy" concept, whose background characteristics represent individual references. The reference code (e.g. ***ELSHAMI90***) is represented by the characteristic name, while the full bibliographic details (e.g. *Elshami, Ahmed M. 1990. <CD-ROM Technology for Information*

[35] Concept clusters are groups of related concepts which are somehow related. Their relations have not yet been made explicit because the terminologist has lacked either the time or the knowledge to do so.

Managers.>[36] *Chicago: American Libraries Association, 387 pp.*) make up the characteristic value.

3.2.4 Establishing a Concept System

From further reading of subfield documentation, the terminologist should be able to identify what appear to be the most common concepts in that subfield and establish a preliminary concept system. The graph allows the terminologist to easily explore concept relations from the beginning. For example, as seen in Figure 24, our initial subfield graph showed that the subfield of optical storage media was divided up hierarchically, with concepts having generic-specific relations with each other. In fact, we found that there were three main "kinds" of optical storage media: read-only, write-once, and erasable, each with its respective specific concepts.

Even at this early stage, CODE's graph enables the terminologist to easily view a given concept's generic concept and any co-concepts.

3.2.5 Enlisting the Help of Subject-field Experts

Since the information gathered at this stage provides the fundamental conceptual structure of the TKB, it is important that the initial generic-specific relations be as accurate as possible. Although there is no specific moment at which to bring in expert help on a terminology research project, this is often the best time.

The help of subject-field experts is essential if the information gathered for the project is to be considered reliable and complete. While terminologists are certainly language experts, they may not know all there is to know about a certain subject field, especially if it is a fairly technical one. CODE's features, including the various browsers with their hierarchical lists of concepts and

[36] Angled brackets are used to indicate italics in COGNITERM, since CODE cannot provided such typographical features.

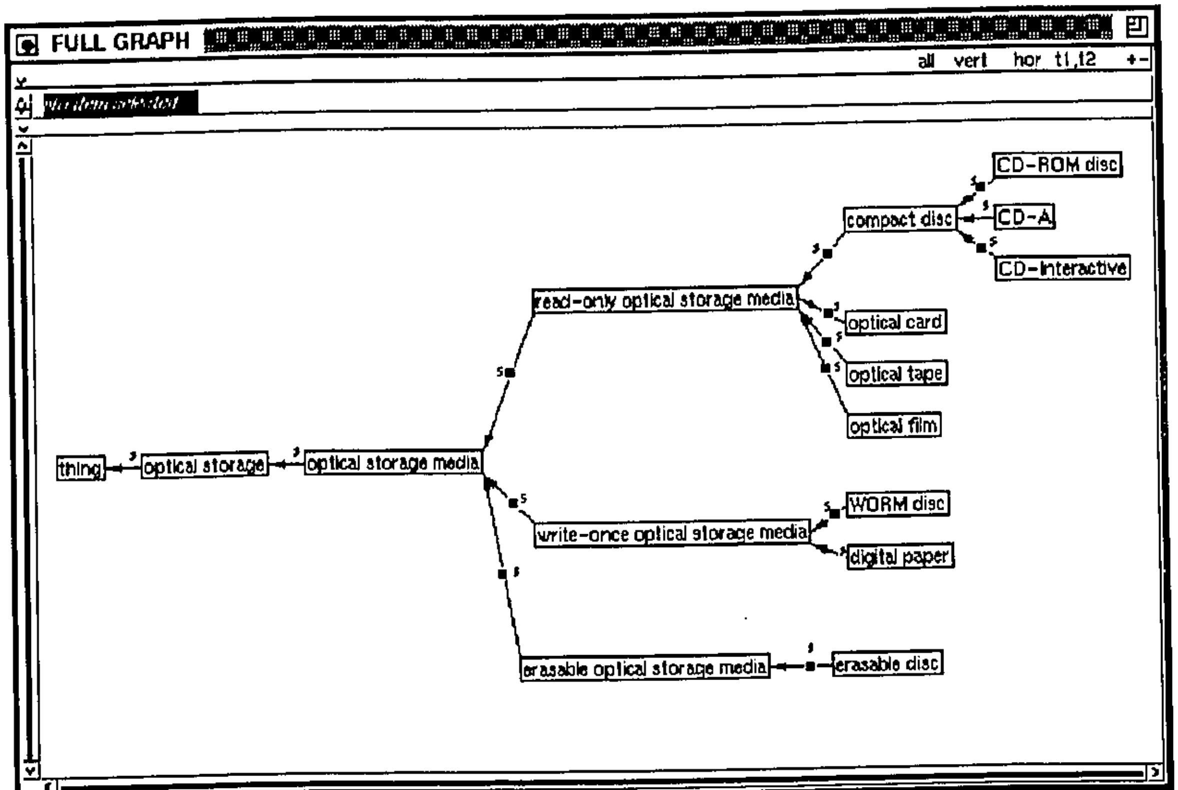


Figure 24. Concept graph for the subfield optical storage media.

characteristics, the graphs, and the Characteristic Comparison Matrix, allow the terminologist to quickly and effectively show potential experts the initial concept system of the selected subfield and to ask informed questions. The subject-field experts may also be able to make suggestions regarding documentation at this point.

3.2.6 Designing a Template of Conceptual Characteristics

The terminologist must also do some preliminary reading of the primary documentation to begin designing a TEMPLATE of the conceptual characteristics that appear to be common to the concepts (common in that they share the same CN). While this is not an easy task, it is an extremely important step, since it is these characteristics which the terminologist eventually uses to distinguish one concept from another.

Chapter 1 (section 1.2.3.2.1) explained that several kinds of conceptual characteristic must be considered. COGNITERM provides a default template for the convenience of the terminologist, as seen in Figure 25. In this template, ***Conceptual Characteristics*** are divided up into two main categories: ***Attributes*** and ***Relations***, with ***Relations*** further divided into ***Associative relations*** and ***Partitive relations***. The terminologist will have to decide which characteristics specific to the concepts being explored should be added to this template, and whether they are relations or attributes.

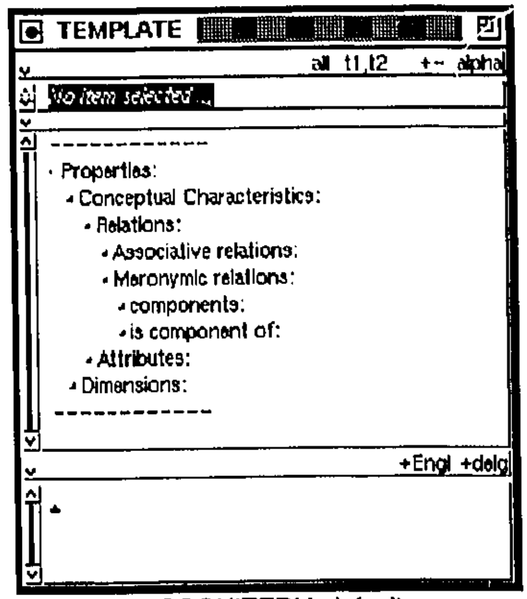


Figure 25. COGNITERM default conceptual characteristics template.

It is with this template that the terminologist can ensure that the conceptual information remains consistent and explicit. The subfield documentation provides information in the form of contexts, which present the conceptual information implicitly. When this information is converted into a characteristic, it is given a characteristic name (e.g. ***diameter***) and a characteristic value (e.g. *12 mm*). Any other information is not considered useful to the terminologist (and the end user) is eliminated.

Table 3 shows how contexts found for the concepts **CD-Interactive** and **CD-ROM** each yielded four different characteristics (even more information could have been made available for the concept **CD-I drive**).

With such rich contexts, there is a danger of the terminologist becoming confused when trying to compare and contrast concepts and determine what exactly differentiates them. When the information is presently explicitly in the form of characteristics, the terminologist's task of comparing concepts is much easier.

Context	*Characteristic Name:*	*Characteristic Value*
Compact Disc Interactive is a specific set of standards for storing multimedia information (text, audio, graphics, video) on a Compact Disc. It was developed by Philips and Sony, and is defined in the Green Book. CD-I discs can only be "played" on special CD-I drives, which include their own microprocessor (Motorola 68000 series) and use their own operating system (CDRTOS or Compact Disc Real Time Operating System).	***content:***	*text, audio, graphics, video*
	standard:	*Green Book*
	drive used:	*CD-I drive*
	developers:	*Philips and Sony*
CD-ROM is a set of standards for storing 600 Megabytes and above of digital information on a Compact Disc. It was developed by Philips and Sony, and is defined in the Yellow Book.	***capacity:***	*600+ Megabytes*
	content:	*digital information*
	developers:	*Philips and Sony*
	standard:	*Yellow Book*

Table 3. Contexts and the characteristics derived from them.

Although there is no strict methodology for selecting the main characteristics for the template, several "rules of thumb" can be followed:

(1) The terminologist should note characteristics that repeatedly appear in the documentation. For example, we found that the contexts in Table 3 were a good starting place, given that several of the characteristic names appeared in each context.

(2) Characteristics which appear in "defining contexts" (i.e. contexts which appear to give a good overview of the concept) are worth noting.

(3) Characteristics which are inherited by each co-concept from the same generic concept, but whose values change should be recorded.

(4) Lists or tables of information are sometimes found in documentation, and these can often be adapted by the terminologist as characteristics. For example, several documents on optical storage media included "disc specifications"; these concise lists included such characteristic names as ***diameter***, ***storage capacity***, ***error correction***, ***material***, and ***standard*** (along with the appropriate values), which we added directly into our template.

The terminologist may find it helpful to follow some simple conventions devised by the COGNITERM team regarding the style and syntax of characteristic names and values:

— if a term already exists to express the characteristic name or value, it should be used; for example, in the optical storage TKB, the characteristic names ***physical standard*** and ***logical standard*** are both found in a number of documents on optical storage technology

— if no term exists to express the characteristic name or value, the terminologist should create one that is as transparent as possible, ideally in consultation with an expert; for example, in the optical storage TKB, the characteristic name ***degree of writability*** was created by the COGNITERM terminologists to describe the values of *read-only*, *write-once*, and *erasable*, since no documented characteristic name could be found. This name was eventually approved by an optical storage expert.

— characteristic names should be nouns or noun phrases as often as possible to make them easy to refer to in discussions among terminologists, and between terminologists and experts. For example, the terminologist might want to ask questions of the expert such as "What do you know about the CD-ROM's ***error correction code***?" or "Is the ***error correction code*** X or Y?"

— characteristic values should be as simple and succinct as possible (i.e. not an entire context, but the important information contained in the context, as demonstrated in Table 3)

— a generic concept may have a characteristic value that in effect constitutes an amalgamation of the characteristic values of its specific concepts; if this is the case, the value should be expressed in the following syntax:

— *one of:<list>*, where <list> is a list of all the characteristics present

e.g. the concept **storage** has the characteristic ***optical storage technology:one of:optical, magnetic, magneto-optical***; each of its specific concepts can have only one of these values

— *one or more of:<list>*, where <list> is a list of all the characteristics present

e.g. the concept **compact disc** has the characteristic ***content:one or more of:textual, audio, video***; each of its specific concepts can have one or more of these values

The subject–field experts should also be consulted once a preliminary template has been constructed. Their input at this stage is invaluable, since the template influences the subsequent work of the terminologist.

It must be remembered that this template is not carved in stone. It may well require modification at later stages, and will certainly have new characteristics added at lower levels in the hierarchy as new concepts are added. However, a top–level characteristics template that is well–designed from the beginning will make it easier for the terminologist to begin to gather the conceptual information she needs.

3.2.7 Scanning

The next major step in concept analysis is scanning, which involves careful reading of the primary documentation and extraction of both conceptual and linguistic information.

All of this information is added directly to the TKB via the various browsers. As noted in sections 3.1.2 and 3.1.3, information about the concept is added to the Conceptual Info browser as conceptual characteristics, while information about the term is added to the Linguistic Info browser as linguistic characteristics. All characteristics must be fully referenced, with a reference code accompanied by a page number provided in the ***REFERENCE*** facet.

Linguistic characteristics can be added to the linguistic characteristics template in the

Linguistic Info browser (synonyms, phraseological contexts, defining contexts, and derivates). For conceptual information, the terminologist analyzes contexts found in the documentation and adds any pertinent value to the appropriate characteristic, following the conceptual characteristics template created earlier. The template may be altered if necessary: if other characteristics are found which appear to be crucial to understanding the concept, they should be added, ideally after consultation with an expert.

3.2.8 Revising the Information

Before the project is completed, the collected information should be reviewed once again by subject-field experts. The terminologist should ensure that she takes full advantage of CODE's features to ensure that the information is accurately presented to the revisers. Well-structured, explicit information not only speeds up the verification process, it also facilitates any necessary modifications.

3.2.9 Updating the information

Once the terminology research project is concluded, periodic updating of records should be done where warranted, since it is possible that the information about the concepts and their terms may evolve, especially in more modern, technical fields. Fortunately, it is always possible to revise and update this information, since CODE allows the terminologist to keep "dynamic" records. For example, although we had completed our research on optical storage media, further information was found on **CD-ROM XA** and **DVI** during the scanning of documentation on optical storage devices. This information was easily added via the browsers. As well, several new concepts "appeared" during this time (i.e. new products were introduced onto the market), including **Kodak Photo-CD**, **CD-G**, and **CD-R**. They too were easily incorporated into the TKB.

3.3 CONCLUSION

Chapter 2 described CODE's strength as providing the terminologist with a more methodical way of managing the information being collected than pencil-and-paper or do-it-in-your-head methods. The information is presented in a structured and explicit form, and can be easily added to and modified.

This structure and explicitness facilitates the construction of intensional terminological definitions. The various aspects of definition construction within a knowledge-based framework are the discussed in part 2.

Part 1 has introduced the basic concepts, technology, and approach with which the terminologist must be familiar in order to carry out general concept analysis using CODE, a knowledge management tool. The main points of this introduction are the following:

(a) The terminologist must have a thorough grasp of what describes a concept, i.e. characteristics (attributes and relations), in order to be able to carry out general concept analysis.

(b) As a general knowledge management system with numerous special features, CODE facilitates the general concept analysis process, and allows the terminologist to undertake this task in a much more methodical way than was previously possible with pencil-and-paper and "do-it-in-your-head" methods.

(c) By following the approach outlined in chapter 3, which in effect comprises the basic steps to be followed for general concept analysis, the terminologist can effectively use CODE to record and display the results of her concept analysis.

PART 2

DEFINITION-ORIENTED CONCEPT ANALYSIS: THE ITD, THE TECHNOLOGY, AND THE APPROACH

Part 2 of this thesis discusses the changes proposed both to the technology and to the approach presented in part 1 in order to support a definition-oriented approach to concept analysis.

First, Chapter 4 surveys what is found in the terminology literature (and, to a lesser degree, some of the lexicography literature) about intensional definitions. The results provide the basis for the proposed refinements and additions. The intensional terminological definition (ITD) is explored from several points of view: (1) as a product, i.e. the components of the ITD; (2) as a process, i.e. how the terminologist should go about constructing this product; (3) and as a tool, i.e. the roles that the ITD fulfils.

Chapter 5 then explores the design and development of the Characteristic Comparison Matrix (CCM), a CODE feature of particular relevance to definition construction. The CCM was introduced in the previous version of CODE (CODE2) and underwent extensive testing and analysis by me before being implemented in CODE4.

Finally, Chapter 6 suggests how the general approach to concept analysis (outlined in chapter 3 of this thesis) can be enhanced to allow the terminologist to construct ITDs upon the completion of concept analysis. While the actual writing of definitions might be considered a "post-concept-analysis" activity, the resulting definitions are obviously based on the information gathered during concept analysis; thus the terminologist must start to prepare for the task of definition construction long before concept analysis actually comes to an end. The proposed enhancements consist of both refinements and additions to the existing general approach. The resulting approach is termed *definition-oriented concept analysis*.

Chapter 4 THE INTENSIONAL TERMINOLOGICAL DEFINITION

4.0 INTRODUCTION

The introduction to this thesis described how definitions have been studied for over two thousand years. Philosophers and linguists who have shown interest in them range from Aristotle to Thomas Hobbes to Eugen Wüster, the founder of the discipline of terminology. A survey of the literature might give the impression that there are almost as many kinds of definition as there are definition scholars, as suggested by the various definition typologies noted in section 0.2.2. For the purposes of this thesis, however, I am concentrating on the intensional terminological definition (ITD).

This chapter explores the intensional terminological definition from several different perspectives: as a product (the ITD and its components); as a process (how the ITD is constructed); and as a tool (what the ITD can be used for). Process and product are closely linked, since the way in which the process is carried out influences the quality of the resulting product. As a tool, the ITD fulfils several different roles, providing the user with information about the concept in a variety of ways.

What is most striking in a survey of the literature is the lack of detail on the actual process of defining. Hints are provided on how the terminologist can carry out the process, and more specific information is provided about the resulting product, as well as its role. But nowhere, unfortunately, is there an easy recipe to follow which produces perfect ITDs every time. And with good reason: despite the claim of *Webster's Ninth New Collegiate Dictionary* that "orismology [is] the science of defining technical terms," the consensus of terminology researchers is that defining is more art than science, with the terminologist carefully weaving together the various elements of the ITD by combining knowledge, intuition, and creativity. If the terminologist works methodically, collects the various components of the ITD with care, applies certain rules of thumb, and verifies her work with subject-field experts at different stages, she stands a considerable chance of creating a good ITD.

4.1 THE ITD AS PRODUCT

The intensional definition has always been considered an essential part of terminology work, regardless of the technology used (i.e. pencil-and-paper or sophisticated computerized systems) or how the ITD was generated (i.e. extracted from contexts or composed by the terminologist). Because of its origins,[37] it is variously described as "traditional" [Béjoint 1988:364], "age-old" [Dahlberg 1983:22], and "classical" [Picht/Draskau 1985:51]. The intensional definition assembles the most important information about the concept in order to give the user an immediate understanding of its essential features. It expresses (a) the relationship between the concept being defined and its generic concept and (b) the characteristics which make it more specific than its generic concept and differentiate it from its co-concepts, all in a clear, concise package.

As Picht/Draskau [1985:51] note, the intensional definition is like a mathematical equation, in which the two halves are:

- the ***definiendum***: the term designating the concept being defined
- the ***definiens***: a combination of the GENUS (the term designating the generic concept of the concept being defined) and the DIFFERENTIA (the terms designating the necessary and sufficient (n+s) characteristics of the concept being defined).

Figure 26 shows the parts of the ITD, along with examples.

[37] As noted in the introduction (section 0.3.1), the intensional definition was devised by Aristotle, nearly 2500 years ago.

ITD:	**definiendum**	=	**definiens**		
			genus	(+)	differentia
	⇓		⇓		⇓
	term designating the concept being defined	=	term designating the generic concept of the concept being defined		terms designating the n+s characteristics of the concept being defined
example 1	A CD-ROM	is a	compact disc		that is read-only and can store up to 600 MB of textual data.
example 2	A CD-Audio	is a	compact disc		that is read-only and can store up to 74 minutes of audio data.

Figure 26. The formal structure of the intensional terminological definition (adapted from Picht/Draskau 1985:51).

Three things should be noted about this structure. First, the definiendum should be synonymous with the definiens, i.e. if the definiens is used in a sentence in place of the definiendum, the meaning of the phrase should not change.[38] A comparison of two sentences, one containing the definiendum of the ITD of **CD-Audio**, the other the definiens, illustrates this.

A sentence using the definiendum would read:

> I don't like any of the songs on the *CD-Audio* that I bought yesterday.

A sentence using the definiens would read:

> I don't like any of the songs on the *compact disc that is read-only and can store up to 74 minutes of audio data* that I bought yesterday.

Second, it should be noted that, in an intensional definition, the generic-specific relation between the concept being defined and its generic concept is made explicit. For example, the concept **CD-ROM** has a generic-specific relationship with the concept **compact disc**, where compact disc is the generic concept: i.e. a **CD-ROM** ***is a*** kind of **compact disc**. It is important that the generic-

[38] While the meaning of the phrase should not change, the style may become unnatural or awkward.

specific relation be the same for all co-concepts, a fact which is discussed further in section 4.1.1.

Third, the concept being defined is shown to be more specific than the generic concept, and differentiated from its co-concepts, by means of its n+s characteristics. While it is not easy to decide which characteristics differentiate a concept from its co-concepts, there are some rules of thumb which the terminologist can follow. These rules are discussed further in section 4.1.2.

4.1.1 The Genus: Consistent Generic-Specific Relations

Section 1 explained how the primary relation represented by the ITD is the generic-specific relation. Ten Pas [1991:79] suggests that the first step in ITD construction is the classification of related concepts into a concept system. Such a system would show the generic-specific (and perhaps other) relations that exist between the concepts, and help the terminologist determine the genus for the ITD.

For the ITD to correctly provide the user with an understanding of the relations among concepts in a particular subject field,[39] it is important that the ITDs of co-concepts show that each co-concept has the same generic concept (denoted by the genus). While this may seem obvious, a small study of "coordinate" ITDs shows that problems with inconsistent generic concepts occur frequently.

The following example, combining Table 4 and Figure 27, compares a graphical display of part of the COGNITERM Typesetting TKB with ITDs found in high quality records from the TERMIUM III term bank.

[39] See section 2 for further discussion of the roles of the ITD.

Concept	COGNITERM graph	TERMIUM III definition	Comments
laser printer	laser printer →[40] non-impact printer → printer	A **nonimpact printer** that operates at well over 10,000 lines per minute, using a low-power laser to produce image-forming charges a line at a time on the photoconductive surface of a drum...	Although the COGNITERM graph shows these to be co-concepts, TERMIUM III gives them different generic concepts (i.e. **nonimpact printer** *vs.* **nonimpact printing device**; printing device is in fact the generic concept of printer).
thermal printer	thermal printer → non-impact printer → printer	A **nonimpact printing device** in which a special, heat-sensitive paper is passed over a dot matrix of heating elements...	
Postscript printer	Postscript printer → laser printer → non-impact printer → printer	A **printer** using the PostScript page description language.	The generic concept provided by TERMIUM III (i.e. **printer**) is three levels too high.
daisywheel printer	daisywheel printer → impact printer → printer	A type of **serial impact printer**[41] in which the font is formed on the end of spring fingers that extend radially from a central hub.	These are shown to be co-concepts on COGNITERM graph, but this cannot be seen in TERMIUM III definitions.
dot matrix printer	dot matrix printer → impact printer → printer	A **printer** that prints characters or images represented by dots.	

Table 4. A study of generic concepts (in bold) within definitions in TERMIUM III.

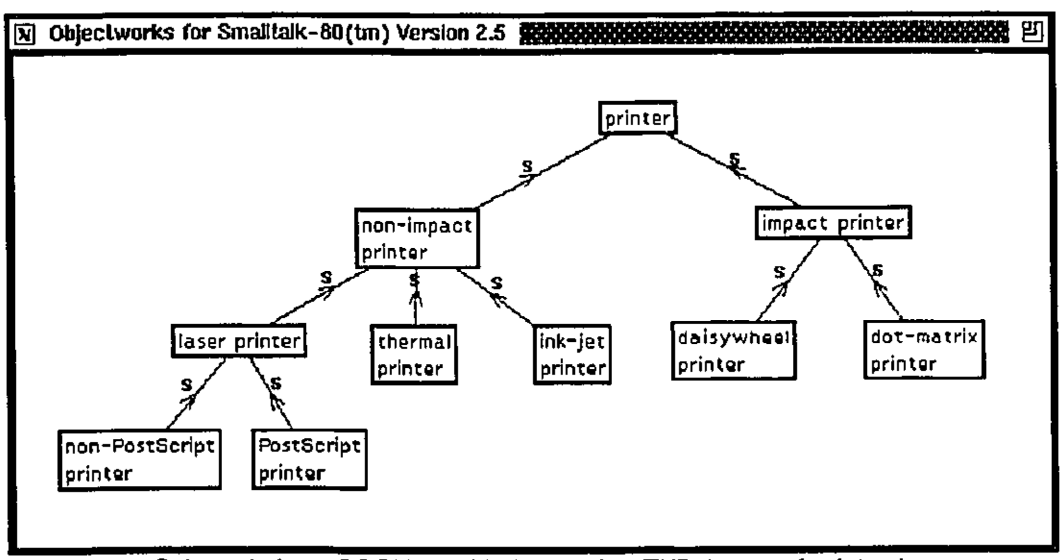


Figure 27. Subgraph from COGNITERM Typesetting TKB (types of printers).

[40] The arrow (→) can be read as "is a" or "is a kind of." For example, "a **laser printer** is a kind of **non-impact printer.**"

[41] It may be that **serial impact printer** is a concept (TERMIUM III marks "unconfirmed" in a B-quality record), and that COGNITERM is missing a level of concepts.

The graph (which explicitly shows the generic–specific concept relations within the subject field) is one way of demonstrating how constant reference to a clearly marked hierarchy enables the terminologist to construct intensional definitions that have consistent generic concepts.

As the comments in Table 4 indicate, there is a lack of consistency in the TERMIUM III definitions regarding generic concepts. This may be due to the fact that the TERMIUM III method of definition construction is primarily context-based[42]: contexts hardly ever consistently provide the same generic concept for all coordinate concepts. When ITDs are constructed using a concept system that clearly displays the relationships between the various concepts (Figure 27), there is less danger of such inconsistency.

4.1.2 The Differentia: Necessary and Sufficient Characteristics

Characteristics are very important to ITD construction, as all specialists note [cf. Picht/Draskau 1985, Sager 1990, Dubuc 1985]. The sum of the characteristics (i.e. the intension of a concept) is equal to the sum of our knowledge about a concept, according to Picht/Draskau [1985:40]. While a concept's intension may include many characteristics, only a subset of these, the n+s characteristics discussed above, is typically included in the intensional definition. Dahlberg [1987:7] notes that definitions "are the shortest form of a [concept's] description. They contain just those elements which are absolutely necessary for its characterization...." While there are those who espouse the need for "encyclopedic" definitions, which contain as much information about the concept as possible (e.g. for non-specialists requiring "some degree of understanding of the term" [Sager 1990:49, Dahlberg 1981:17]), de Bessé [1992:183] counters that:

[42] According to TLS [1984:3–6], three types of definitions can be used for TERMIUM: handcrafted definitions (constructed by a qualified terminologist, group of terminologists or committee); composite definitions (composed of bits and pieces taken from definitions found in other sources) and quoted definitions (taken verbatim from the source). The emphasis appears to be on definition selection, however, not on definition construction.

> La définition terminologique n'exprime pas nécessairement tous les traits pertinents d'une notion. La définition peut se limiter aux éléments nécessaires et suffisants à la compréhension du sens du terme par les usagers [....] Il n'est pas indispensable qu'elle donne une description exhaustive de la notion définie.

Two questions come to mind here. First, what *are* "necessary and sufficient" (n+s) characteristics? Second, how does one decide which characteristics fall into this category? The answers to these questions are explored below.

4.1.2.1 Necessary Characteristics

For a characteristic (C) to be necessary to a concept (X), one must be able to say that "X must have/be C". For example, some necessary characteristics of the concept **river**, whose generic concept is **body of water**, are:

1. *origin:natural*
2. *motion:flowing*
3. *width:>W*[43]

If a concept was missing one of these characteristics, it would simply not be a **river**. For example:

a. if it were not natural, it might be a **canal**;
b. if it didn't flow, it might be a **lake**;
c. if it had a width <W, it might be a **stream.**

It should be noted that what is necessary in one language may not correspond to what is necessary in another. The well-known distinction in French between the concepts of **fleuve** and **rivière** support this point. The principal characteristics which differentiate these concepts are the river's width, and what it flows into, which could be represented as *flows into:one of:sea or larger river*. These characteristics in fact dictate whether the English **river** is translated as **fleuve** or **rivière.**[44] While it is extremely important to include these characteristics when describing **fleuve** (i.e.

[43] W represents a certain width which the terminologist must determine.

[44] Although this analysis treats the two concepts as co-concepts, some dictionaries (e.g. *Petit Robert*) consider **fleuve** to be a kind of **rivière**.

width:>W, flows into:sea) or **rivière** (i.e. *width:<W, flows into:larger river*), such is not the case for the English **river**. Thus, while **river** has at least three necessary characteristics, both **fleuve** and **rivière** have at least four.

There are other characteristics used to describe a concept which are nevertheless not necessary; they are considered typical,[45] but not needed for a basic understanding of the concept, since not every single instance of the concept has these characteristics (i.e. not every instance of X must have/be C). Two examples of typical characteristics for the concept **river** are:

4. *fauna:fish*
5. *application:navigation*

While **rivers** *may* have fish and be used for navigating, a **river** without fish is still a **river**, as is a **river** that is not used for navigational purposes. These characteristics are not necessary for a basic description of the concept.

4.1.2.2 Sufficient Characteristics

Sufficient characteristics are a set of characteristics that meet a special criterion: for a set of characteristics {S} to be sufficient for a concept (X), one must be able to say that {S} implies X; i.e. something that has {S} is automatically represented by X. Mel'čuk [1988:31] notes that the set of N+S CHARACTERISTICS and the concept "doivent être mutuellement substituables...dans tous les contextes imaginables...." For example, the three necessary characteristics provided for **river** in section 4.1.2.1 are considered to form the sufficient set.

[45] In some cases, a typical characteristic is so important that it should be included in a definition. For example, the concept **bird** has the characteristic *mode of transportation:flying*; this is a typical characteristic, since not all **birds** fly (e.g. **ostrich, penguin**). However, it could be considered a candidate for inclusion in a definition, since it is one of the main characteristics which distinguish **birds** from other animals.

4.1.2.3 Necessary and Sufficient Characteristics

The proper combination of necessary *and* sufficient characteristics ensures that the concept described by the characteristics is neither too generic, nor too specific. For example, the set of characteristics (1–3) in the previous section is sufficient to describe a **river**, given the genus of **body of water**. However, one could imagine a set of necessary characteristics that, taken together, are not sufficient. In that case, the set of characteristics would be describing something more generic than X, since one or more n+s characteristics have been left out. For example, the following two characteristics are necessary to a **river**:

1. *origin:natural*
2. *motion:flowing*

In other words, all **rivers** must be *natural* and all **rivers** must *flow*. However, a concept with only these two characteristics (and not the third, *width:>W*) could also be a **brook**, or a **spring runoff**. Each of these concepts is *natural* and *flows*. Thus, this particular set of characteristics is not sufficient to define a **river**: a definition containing only these n+s characteristics would be too broad.

There could also be a set of characteristics thought to be sufficient to define X, but one of the characteristics is not necessary; this set of supposedly n+s characteristics describes a concept that is more specific than X. For example, if the characteristic *application:navigation* is added to the ITD of **river**, the ITD becomes too narrow, as it would exclude some **bodies of water** considered **rivers** that are not, in fact, used for navigation.

Thus, it is the n+s characteristics (denoted by the term "differentia") which differentiate the concept from its generic concept, as well as from its co-concepts. While co-concepts can share some of the same n+s characteristics, each concept has at least one n+s characteristic which differentiates it from the others.

	Characteristic Values			
Characteristic Name	**River**	**Canal**	**Stream**	**Lake**
origin:	*natural*	*man-made*	*natural*	*natural*
motion:	*flowing*	*flowing*	*flowing*	*standing*
width:	*>W*	*>W*	*<W*	*>W*

Table 5. The characteristics of different kinds of **body of water.**

For example, as Table 5 shows, **river** is differentiated from other kinds of **body of water** by the nature of its ***origin***, ***motion***, and ***width***. Thus, these characteristics are included in a possible ITD of **river**:

A **river** is a *natural*, *flowing* **body of water** with a *width* >*W*.

A canal may have the same characteristics as river except for its ***origin*** (*man-made*). Thus its ITD could be:

A **canal** is a *man-made*, *flowing* **body of water** with a *width* >*W*.

A **stream**, however, while a co-concept of **river** and **canal**, differs from them both by being narrower. Thus its ITD could be:

A **stream** is a *natural*, *flowing* **body of water** with a *width* <*W*.

Lastly, a **lake** could be defined in the following way:

A **lake** is a *natural*, *standing* **body of water** with a *width* >*W*.

4.2 THE ITD AS PROCESS

While it is one thing to describe the definition as a combination of genus and differentia, it is another to actually isolate these components from the available information about a concept. Various aspects of the genus and differentia selection are described below, with particular emphasis on the latter.

4.2.1 Selecting the Genus

It is not always an easy task to decide upon the genus of an intensional definition, and in fact, Rey–Debove [1966:79] feels that "[c]ette recherche du genre prochain est la difficulté majeure de la définition." The problem lies in the fact that the concept to be defined may often have more than one generic concept. There appear to be three main reasons for this: (1) multidimensionality; (2) the division of concepts according to a pre–established ontology[46] vs. the division of concepts according to subject field; and (3) insufficiently classified subfields.

Multidimensionality

Bowker 1992 explains how the concepts of a given subject field can be classified more than one way, depending on the classifying characteristic (the characteristic used to divide up the subject field). Each classification is called a dimension, and a subject field can have (and usually *does* have) more than one dimension. This phenomenon is called multidimensionality.[47]

For example, Bowker [1992:17] shows how the concept **vehicle** can be classified according to the following dimensions:

— by ***medium of transportation***: **land vehicle, air vehicle, water vehicle**

— by ***type of propulsion***: **motorized vehicle, non–motorized vehicle**

— by ***type of load***: **passenger vehicle, cargo vehicle, passenger and cargo vehicle**

Figure 28 shows that the specific concepts of the above concepts (e.g. **land vehicle, air vehicle**) have more than one generic concept if all dimensions are considered. The terminologist

[46] An ontology can be loosely defined as the selected set of "very general semantic categories" [Skuce 1991a:3].

[47] Bowker [1992:20–28] suggests several reasons why a subject field may be classified according to different dimensions, including viewpoint, purpose of classification, opinion, scientific theory, and language and culture.

constructing definitions within one dimension[48] would have to decide whether an **airplane** is an **air vehicle**, a **motorized vehicle**, or a **passenger and cargo vehicle**. This decision will likely have to be based on an expert's opinion as to which dimension is the most recognized in the field.

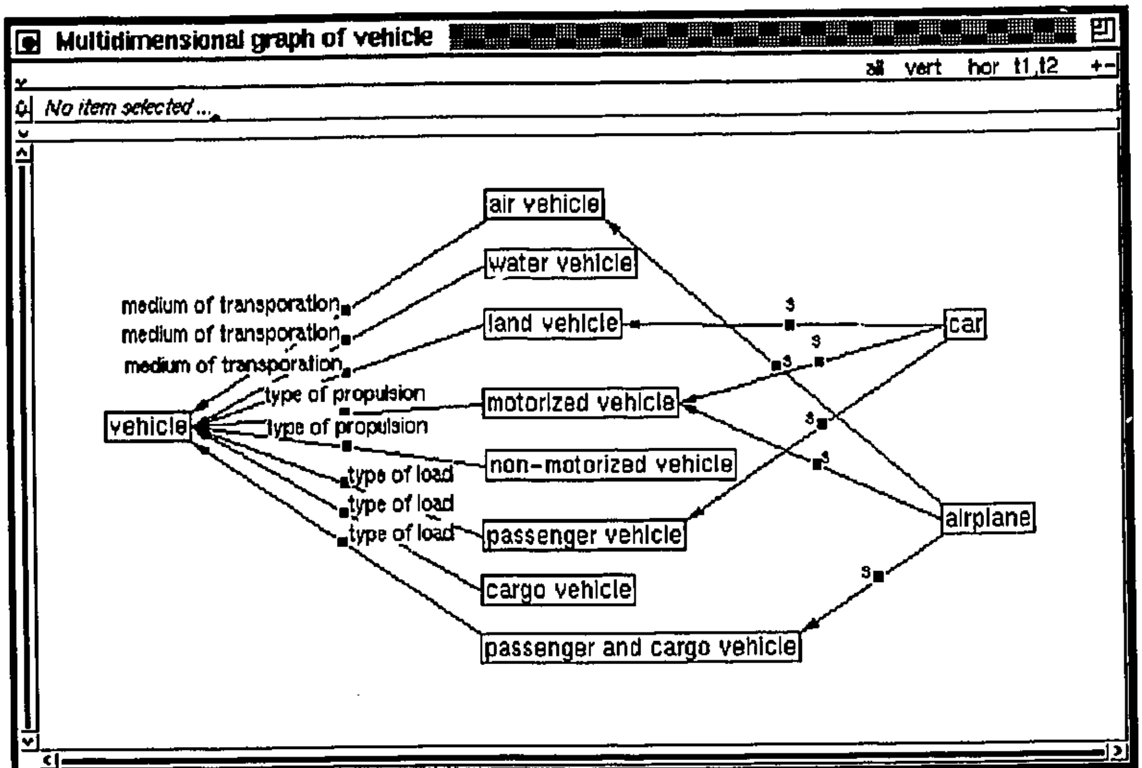


Figure 28. A multidimensional classification of **vehicle** (adapted from Bowker 1992:15).

Ontology vs. subject field

The generic concept of a given concept can also vary according to whether the terminologist follows a general ontology or lets the subject field dictate general concept relations. For example, Pearson [1992] uses an ontology of general categories such as **function**, **equipment**, and **process**, so that every concept is a kind of **function**, a kind of **equipment**, etc. However, Pearson also considers every concept to have a generic concept that is directly "above it" in the hierarchy of a specific subject field, so that, for example, a **television signal** is a **signal** [Pearson 1992]. Thus, a given concept

---

[48] In principle, the terminologist could construct a different definition for each dimension. In this thesis, however, I will focus on what appears to be the most important dimension so that, in the end, only one definition is constructed.

actually has two generic concepts: one stemming from a general ontology, and the other from the subject field hierarchy.

In the optical storage TKB, we initially chose to let the subject field dictate the main generic-specific relations, so that, for example, **CD-ROM drive** is a kind of **optical storage device**, and **optical disc** is a kind of **optical storage media**. It is anticipated, however, that within COGNITERM, the TKB will eventually be linked to a general ontology, so that each concept has two generic concepts: a general one based on an ontology, and a specialized one based on the subject field hierarchy.[49]

This problem of which approach to take when selecting the genus of a concept-to-be-defined appeared when I constructed a small TKB[50] on the **performance specifications** of optical storage devices, complete with definitions. (While the definitions were essentially modified defining contexts, an effort was made to have consistent genuses among co-concepts.) Although I felt that the genus for the concepts in the TKB (e.g. **seek time, access time, transfer rate**) should be **performance specification,** my project supervisor (Dr. Geneviève Mareschal) felt that the genus of **seek time** and **access time** should be **time,** and that of **transfer rate** should be **rate**. I would submit that we each took a different approach to genus selection: I had opted for the more specific, subject-field approach, while Dr. Mareschal had chosen the more general, ontological approach. Pearson [1992] and the COGNITERM Project indicate that perhaps both approaches should be used to show that each concept may have more than one genus.

---

[49] The ontological generic concept will be at a much higher level and *not* an immediate generic concept, as the term "generic" suggests. For example, the top-level concept of both CD-ROM drive and optical disc might be "physical object." See Miller 1993 (section 4.3.3) for more on the inclusion of high-level concepts in a TKB.

[50] This TKB was constructed as part of a terminology project during my Master's course work.

Insufficiently classified subfield

It may also be that the terminologist has insufficiently broken down a subfield, so that what might appear to be the correct generic concept is in fact a "super" generic concept. For example, as seen in Figure 29, the concept **body of water** (discussed in section 1.2) could have been further divided into **flowing body of water** and **standing body of water**, according to the characteristic *motion:one of flowing or standing*. Since only one concept with the characteristic *motion:standing* was considered in the example above (i.e. **lake**), the importance of this characteristic was not

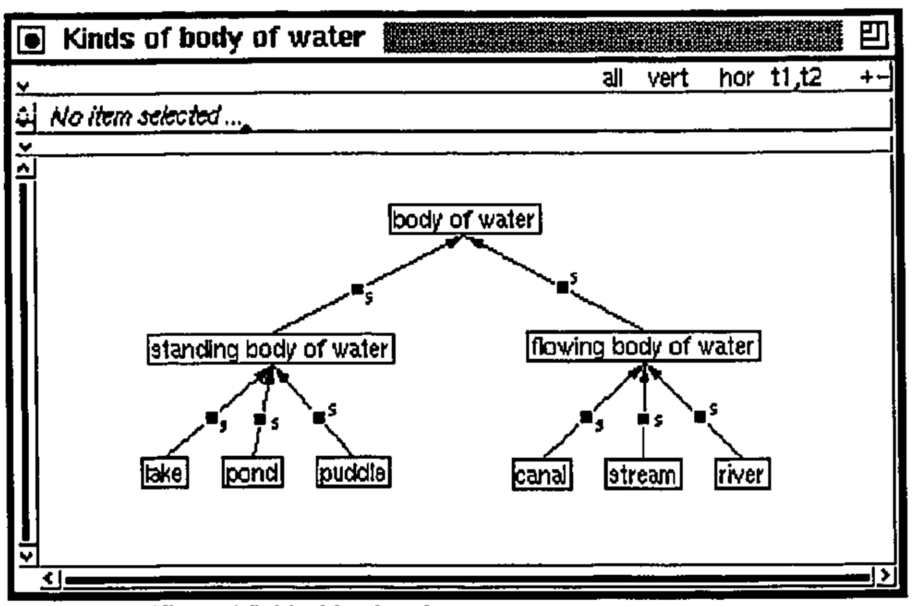


Figure 29. The subfield of **body of water.**

immediately noticed. However, when more concepts with the characteristic *motion:standing* are added (e.g. **pond** or **puddle**), the need to further divide the concept **body of water** becomes more obvious. The resulting definitions for the concepts at the lowest hierarchical level of Figure 29 should thus have either **flowing body of water** or **standing body of water** as their genus.

To summarize, the terminologist must take several factors into consideration when classifying the subject field and selecting the genus for the definitions. These include dealing with multidimensionality, choosing between a general ontology and a more specific subject-field classification, and ensuring that a subfield is divided up according to the n+s characteristics. In the end, it may be the experts who decide on the best way to classify the subject field. What is most important is that the genus of ITDs of co-concepts be *the same*, so that users understand which concepts share the same generic concept.

4.2.2 Selecting N+S Characteristics

Along with the selection of the genus, it is the selection of the n+s characteristics upon which the construction of the ITD rests. Obviously, the terminologist should first record a general set of characteristics that describe co-concepts by analyzing contexts in the subject-field documentation [Dubuc 1985:98]. Auger/Rousseau [1978:33] note that a definition can be taken directly from the documentation; however, in order to get a sound definition, it is probably better to obtain several "contextual definitions" (i.e. contexts containing characteristics which describe the concept) and then fashion them into a single definition. This practice of modifying existing definitions for the creation of composite definitions is also mentioned in the *Terminologist's Handbook* of the Department of the Secretary of State of Canada [TLS 1984:4].

Once the terminologist has scanned the documentation and collected a general set of characteristics, she must then decide (a) which characteristics are necessary and (b) which of the necessary characteristics appear to be, as a set, sufficient to describe the concept and distinguish it from its co-concepts [Sager 1990:24]. The aim should be to formulate an intensional definition that is both clear and concise [Auger/Rousseau 1978:33], providing as much information as possible within the confines of one sentence.

When deciding which characteristics should be included in the ITD, the terminologist must compare and contrast the characteristics collected for the concept and its co-concepts, and decide which characteristics appear to differentiate them. Once the potential n+s candidates have been identified, the terminologist may want to use them to construct draft ITDs, and then analyze the results. One method of analysis is the application of Rey-Debove's [1966] "tout le défini et rien que le défini" rule (discussed below). If the results are positive, the selected n+s characteristics can be used in the final ITD.

4.2.2.1 "Tout le défini et rien que le défini"

For Rey-Debove [1966:72], a good definition is one that upholds the "definiendum = definiens" equation described in section 4.1 (cf. Figure 26): "la définition est «bonne»...si cette égalité est vérifiée, est juste." Rey-Debove recommends that the old formula "tout le défini et rien que le défini" be used to verify this equality.

Boutayeb [1990:41], who has adapted Rey-Debove's formula, suggests using the formula

$$A = [B + C]_D$$

in which

- A is the definiendum
- B is the genus
- C is the differentia
- D is the definiens

The terminologist should use this formula to verify whether the definition positively answers the following questions:

> **(1) regarding "tout le défini":** ***Are all Ds called A?***
>
> **i.e. Are all concepts which have the generic concept (B) and the set of characteristics (C) denoted by the term (A)?**

> **(2) regarding "rien que le défini":** ***Do all As designate a D?***
>
> **i.e. Do all terms (A) denote a concept which has the generic concept (B) and the set of characteristics (C)?**

If both answers are positive, then the definition can be said to be "good." Boutayeb [1990:42] provides the following sample definitions for which these questions can be asked:[51]

(a) **[optical disc]**$_A$: [[Disc]$_B$ [on which data]$_{C1}$ [are recorded]$_{C2}$ [that are readable by an optical device]$_{C3}$]$_D$.

(b) **[digital optical disc]**$_A$: [[high capacity]$_{C1}$ [disc]$_B$ [which allows the archiving of texts and images]$_{C2}$ [via the recording and reading]$_{C3}$ [of digital data]$_{C4}$ [using a laser]$_{C5}$]$_D$.

Table 6 summarizes the questions that Boutayeb [1990:42–43] asks of the above definitions, in order to verify whether the "tout le défini" and "rien que le défini" criteria have been met.

Definition	Question (1): "tout le défini"?	Question (2): "rien que le défini"?
(a)	Are all [[Discs]$_B$ characterized by C1 + C2 + C3]$_D$ called [optical disc]$_A$?	Do all [optical discs]$_A$ designate [[discs]$_B$ characterized by C1 + C2 + C3]$_D$?
(b)	Are all [[Discs]$_B$ characterized by C1 + C2 + C3 + C4 + C5]$_D$ called [digital optical disc]$_A$?	Do all [digital optical discs]$_A$ designate [[discs]$_B$ characterized by C1 + C2 + C3 + C4 + C5]$_D$?

Table 6. Questions resulting from "tout le défini et rien que le défini."

If the answer to each question is "yes," then the equality between definiendum and definiens has been shown within the definition. The set of characteristics denoted by the differentia is sufficient to distinguish the concept from its co-concepts and adequately describe it; thus the definition is "good."

In the early stages of definition construction, it is possible (and perhaps more usual) that the

[51] The following sample definitions have been translated from the original French. The order of the characteristics has been slightly rearranged to avoid overly Gallic-sounding definitions.

response to either of the "tout le défini" questions may be "no." For example, the following examples show the answers that result when Boutayeb's definitions are reformulated.

Reformulated definition (a), leaving out characteristic C3:

[optical disc]$_A$: [[Disc]$_B$ [on which data]$_{C1}$ [are recorded]$_{C2}$.

Question 1 ("tout le défini"):

Are all [[Discs]$_B$ characterized by C1 + C2]$_D$ called [optical disc]$_A$?

Answer 1:

No: magnetic disk and magneto-optical disk (both of which are non-optical) also fit this description. In other words, the characteristics are not sufficient and the definition is too broad.

Reformulated definition (b), adding an extra characteristic, C5:

[digital optical disc]$_A$: [[high capacity]$_{C1}$ [disc]$_B$ [which allows the archiving of texts, images and sounds]$_{C2}$ [via the recording and reading]$_{C3}$ [of digital data]$_{C4}$ [in textual format overleaved with audio]$_{C5}$ [using a laser]$_{C6}$]$_D$.

Question 1: "tout le défini"?

Are all [[Discs]$_B$ characterized by C1 + C2 + C3 + C4 + C5 + C6]$_D$ called [digital optical disc]$_A$?

Answer 1:

Yes. This definition is not too broad.

Question 2: "rien que le défini"?

Do all [digital optical discs]$_A$ denote [[discs]$_B$ characterized by C1 + C2 + C3 + C4 + C5 + C6]$_D$?

Answer 2

No: only CD-ROM XA denotes [[discs]$_B$ characterized by C1 + C2 + C3 + C4 + C5 + C6]$_D$. This definition is too narrow because of C5 (i.e. in textual format overleaved with audio); C5 should be removed because digital optical discs do not need audio data interleaved with the other data.

As the above examples show, applying Boutayeb's two questions to each definition helps the terminologist (a) ascertain whether the definition properly describes the concept and (b) determine the reasons why it does not *if* it does not.

Question 1 ("tout le défini") tests whether the set of characteristics is sufficient: if the set of characteristics is not sufficient (i.e. too few characteristics are included in the set), the definition fails the test. This indicates that the definition is excessively broad [Picht/Draskau 1985:55] and could be used to define any number of concepts. If this is the case, one or more characteristics must be added to the set to make it sufficient.

Question 2 ("rien que le défini") tests that all characteristics are necessary: if too many characteristics (i.e. some unnecessary ones) are included in the set, the definition fails the test. This indicates that the definition is too narrow [Picht/Draskau 1985:55] and that one or more of the characteristics should be removed.

While the formula and resulting questions presented above are helpful to the definition process, they are by no means foolproof. Defining is rarely easy, as Rey–Debove [1966:73] cautions:

> Ce n'est que par tâtonnements que le lexicographe parvient, à l'aide de cette double formule, «tout le défini, rien que le défini», à trouver la meilleure des équivalences possibles, parmi les concepts élaborés spontanément, et de valeur inégale, que se présentent à lui."

What is true for the lexicographer is also true for the terminologist. It is only by systematically determining a concept's generic concept and deciding which characteristics differentiate it from its co-concepts that the terminologist is able to create a good ITD. As was shown in section 4.1.1, generic concepts must be carefully chosen to ensure that all co–concepts share the same genus; otherwise, definitions of co–concepts appear to have different generic concepts, and the user is misinformed about the relations between co–concepts. The same kind of problem can arise if n+s characteristics are not methodically selected, as is discussed below.

4.2.3 Inconsistent N+S Characteristics

If the terminologist is not careful to systematically collect characteristics for co–concepts, rather than indiscriminately extracting definitions from contexts, she risks having inconsistent n+s characteristics within the definitions of co-concepts. A comparison of the TERMIUM III definitions for the co-concepts **daisywheel printer** and **dot matrix printer** (Table 7) shows that the identified n+s characteristics are quite different.

<table>
<tr><th>Term</th><th>TERMIUM III definition</th><th>n+s characteristics</th><th>Comments</th></tr>
<tr><td>daisywheel printer</td><td>A type of serial impact printer[52] in which the font is formed on the end of spring fingers that extend radially from a central hub.</td><td>· font-forming component: spring fingers that extend radially from a central hub</td><td rowspan="2">(1) Although the graph in Figure 27 shows that these are co-concepts, the definitions do not provide any indication that the concepts share any of the same n+s characteristics (value or name).
(2) The fact that these definitions were extracted from existing glossaries and not constructed may help to explain the lack of consistency between the n+s characteristics.</td></tr>
<tr><td>dot matrix printer</td><td>A printer that prints characters or images represented by dots.</td><td>· printed form: characters or images
· type representation: dots</td></tr>
</table>

Table 7. A study of n+s characteristics within TERMIUM III definitions.

By methodically collecting the characteristics of co-concepts and verifying how they differentiate these concepts, the terminologist increases the likelihood that the resulting definitions properly express what differentiates co-concepts.

4.3 THE ITD AS A TOOL

The ITD fulfils several different roles, including establishing a link between the concept and its term, situating the concept within the concept system, establishing interlingual equivalences, contributing to standardization, and educating the user about the concept.

4.3.1 Establishing a Link Between Concept and Term

The definition is used to establish links between the concept and its term [cf. Picht/Draskau 1985]. Rondeau [1981:84] states that it is important that the definition "[fasse] voir le lien qui la

[52] As noted earlier, it may be that **serial impact printer** is indeed a concept (TERMIUM III has marked it *unconfirmed* in a B-quality file), and that COGNITERM is missing a level of concepts.

rattache à une telle dénomination."

As noted in section 4.2, the intensional definition has two parts: the ***definiendum*** and the ***definiens***. The definiendum is the term designating the concept in question, while the definiens includes its genus and differentia. Again, let us look at the definition for **CD-ROM**:

> A **compact disc** that is *read-only* and *able to store up to 600 MB* of *audio and textual data*.

From this definition, the user can ascertain that an object which is a kind of **compact disc** that is (a) *read-only* and (b) *able to store up to 600 MB* of (c) *audio and textual data* is called a **CD-ROM**.

4.3.2 Situating the Concept Within the Concept System

The definition also helps situate the concept within the concept system [Nedobity 1983:251, quoted in Boutayeb [1990:35]]. This situating has two aspects: how the concept is similar to other, related concepts in the concept system, and how it is different from them. The similarities and differences among concepts are captured via their characteristics: as mentioned in chapter 1 (section 1.2.4), a given concept is similar to its generic concept because it inherits characteristics from it. On the other hand, it is different from its co-concepts because some of its inherited characteristics have been modified, or new characteristics have been added.

The genus of coordinate ITDs tells the user that co-concepts share certain characteristics inherited from their common generic concept. The differentia tells the user how a given concept is more specific than its generic concept and how it differs from its co-concepts.

Rey [1992:43] brings up an interesting point about concepts in still-evolving subject fields, noting that:

> ...si le domaine n'est pas déterminé et construit, ou encore s'il s'agit d'un domaine hétérogène, pragmatique et empirique, la définition terminologique pure est absolument impraticable.

The terminologist should keep in mind that the intensional definition is not feasible for concepts which are still evolving and still "fuzzy," because the conceptual structures themselves are not fixed and intensional definitions cannot be created without clear conceptual structures. While the terminologist can perform concept analysis in order to attain some understanding of an evolving concept, a definition would not be a practical way to "capture" this understanding. Term banks often include non-defining contexts (labelled EX and CONT in TERMIUM III, for example) for such concepts.

4.3.3 Establishing Interlingual Equivalences

Another role of the definition, which is linked both to determining the concept and to situating the concept within the concept system, is that of helping to establish equivalence between languages, as explained in some detail in Wright [1993]. By comparing definitions of two terms (each from a different language) that are proposed as equivalent, the terminologist can verify if each term designates the same concept. For example, while the concept **river** was shown to have three necessary characteristics in section 4.1.2, the concept **rivière** was shown to have four, showing that the two concepts differ.

According to Dahlberg [1981:19], "with a knowledge of the characteristics it is also possible to replace a known term for the referent [concept] in one language by a synonymous term in another language." If these characteristics are found within the confines of a definition, the user's job is that much easier.

This particular definition application is of great importance to translators, as well as to terminologists who must provide or confirm foreign language equivalents. Wüster [1968:2.14] proposes that "[t]erms (and especially technical terms) which are encountered in one language by no means always carry the exact shade of meaning which is borne by their presumed equivalent in another." For this reason, Harris [1983:152] suggests that, from a translator's point of view, "in bilingual records, it is often safer to provide as complete definitions as possible in both languages."

Each definition can describe its respective concept to the user, who can then judge whether the target term is equivalent to the source term.

4.3.4 Contributing to Standardization

Since the definition is used to precisely place a concept within a certain subject field, i.e. to establish (1) its most important characteristics, (2) its relations with other concepts, and (3) the term(s) it is denoted by, it also plays an important role in standardization [cf. Picht/Draskau 1985, Dubuc 1985].

The world-wide information explosion has created an enormous need for unambiguous communication [Rousseau 1991:3]. Rapid advances in the technology of various subject fields, as well as the constant development of new subject fields, have produced an exponential increase in the number of concepts, both in the professional world and in everyday life.

These new concepts all require terms, a phenomenon which results in the ever-increasing need for standardized terminologies, both for intralingual and interlingual communication. As Rousseau [1991:3] points out, "L'activité de normalisation terminologique croissante est devenue nécessaire pour éviter le danger de la babélisation terminologique à laquelle on assiste dans plusieurs spécialités." Ten Pas [1991:79] concurs, explaining that "unambiguous communication between experts is considered to be possible only when the concepts have been clearly defined."

Standardized terminologies are created by standardizing bodies such as the International Organization for Standardization (ISO) and national organizations such as the Canadian Standards Association (CSA). The standards are prepared by the technical committees of these organizations, made up of subject-field and linguistic specialists [Galinski 1991:4, Michaud 1991:6]. Terminological standards can also be created within companies and other organizations to ensure standardized "in-house" terminology.

The definition, used to situate the concept within the subject field, is an integral part of the standardized terminology: it contributes to the terminology's role as a tool for information management and quality control in standardization [Galinski 1991:4], ensuring that communication about concepts is as precise and unambiguous as possible.

The development and use of a standardized terminology within a subject field is in the interests of all concerned parties. A standardized terminology means that subject-field experts, translators, and terminologists communicate with each other more effectively, and that subject-field learners study the subject field with confidence, knowing that the terminology is recognized and used by subject-field experts.

Standards should not be regarded as static, never to be changed once established. Picht/ Draskau [1985:17] emphasize that "standardization does not imply linguistic petrifaction; standards are revised at regular intervals...thus assuring a dynamic updating process." The implication here is that the definition must also be dynamic: it should be capable of evolving along with the concept it describes, if necessary.

4.3.5 Educating the User

This last role is one of the definition's most important: the primary function of the ITD is educational. It serves to explain a concept designated by a given term to the users of term banks and terminological publications [cf. Rondeau 1981, Sager 1990, Hermans 1989].

This relatively simple explanation of the ITD's principal role brings up the more complex questions of (a) who definition users are, and (b) how much information they need from a definition.

4.3.5.1 Users of the Intensional Terminological Definition

Obviously, if the ITD is to be considered a tool, its users must be taken into account. When formulating the definition, the terminologist may also want to consider the level of knowledge of expected users, whether they are technicians or scientists, translators or non-experts [Auger/Rousseau 1978:33]. As Rousseau [1983:40] states, "la définition terminologique n'est pas un exercice isolé et le terminologue ne peut faire abstraction du public visé par la définition qu'il rédige."

The "traditional" users of term banks, and by extension, of definitions, include terminologists, translators, subject-field experts, and non-experts [Sager 1990:49], with translators currently appearing to be a particularly large group, in Canada at least. The number of user types is increasing, however. A survey of several authors [cf. Sager 1990, Meyer/Paradis 1991, Pearson 1992] suggests that user types will expand to include subject-field learners, technical writers, documentation specialists, and even computer systems. However, even this list should not be considered exhaustive: anyone who needs access to a subject-field terminology can be considered a potential term bank user, and hence a potential definition user.

4.3.5.2 The User's Subject-Field Knowledge

The main distinction among various term bank and definition users is the amount of subject-field knowledge that they have. User knowledge may be construed as a continuum, with the lay person at one end with very little subject-field knowledge, the specialist at the other end with a great deal, and other users, with varying degrees of knowledge, between these two extremes. A person with expert knowledge of one subject field may know as little as the lay person about another field, while a translator or terminologist who has been working in a subject field for a long time may be considered very close to an expert.

Thus, it is specifically the user's level of subject-field knowledge that must be considered when definitions are being constructed [cf. Picht/Draskau 1985, de Bessé 1990], as this level of knowledge affects the amount of information that the user expects to receive from a definition. The expert user may be satisfied with a clear, concise, one-line definition which employs specialized terminology, while the non-expert may require more information and more general terminology in order to clearly understand the concept. As Sager [1990:196] notes,

> Terminology databases, like dictionaries, attract users with different levels of subject specialisation. Whereas a scientist will expect a very precise definition of a concept, a translator may prefer a less technical definition and an undergraduate student may be looking for a definition more akin to an encyclopaedia entry.

A definition provided by Rousseau [1983:41] for **agneau**[53] helps illustrate this point:

agneau : **Ovin** ***livré à l'abattoir, n'ayant pas d'incisive de seconde dentition.***

According to Rousseau [1983:41]:

> ...une telle définition, qui se limite à quelques traits distinctifs, n'a visiblement pas pour but de donner une description exhaustive de la notion d'agneau, car, à la simple lecture de cette définition, il faudrait avoir derrière soi une longue fréquentation du mammifère en question pour pouvoir le reconnaître. C'est en réalité l'exemple d'une définition destinée aux spécialistes d'un domaine dont le but est de normaliser et de classer des animaux de boucherie.

Rousseau points out that the above definition is by no means an exhaustive description, but instead limits itself to a few key characteristics (i.e. (1) *livré à l'abattoir*, (2) *n'ayant pas d'incisive de seconde dentition*) that serve to quickly identify the concept for experts in that particular subject field. Those with less than expert knowledge will not be satisfied with such a concise definition: they will want more and perhaps different information. (It is unlikely that a lamb's dentition is of much significance to the lay person.)

While the above discussion implies that several levels of definition ought to be made available in every term bank, it should be kept in mind that there is often only one definition per term (and thus per concept) in a term bank. Sager [1983:123] notes that it may be "difficult to satisfy both the layman and the specialist with a single definition," and suggests that "future lexical data banks may...be provided with several levels of specialization of definitions to meet different user needs"

---

[53] This definition is taken from an AFNOR standard [Rousseau 1983:41].

[Sager 1990:48].

Rather than creating numerous definitions, another way of solving the "one definition/many users" problem may be to consider one definition from different points of view. For the subject-field specialists, a definition provides an immediate understanding (or reminder) of the concept. For a subject-field learner, the definition furnishes an introduction to the concept. For those with different levels of knowledge, the definition gives varying degrees of reminder and introduction. For example, one possible definition for **CD-ROM** is:

> A **compact disc** that is *read-only* and *able to store up to 600 MB* of *audio and textual data*.

A specialist would see (or be reminded) that a **CD-ROM** was (a) a **compact disc** that was (b) *read-only* and (c) *able to store up to 600 MB* of (d) *audio and textual data*. Each of these four elements holds a substantial amount of information for subject-field experts. For non-experts, however, they merely represent starting points: they provide a general understanding of the concept **CD-ROM**, and suggest that users who want to know more should explore the concepts found in the characteristics (both names and values). Ideally, the term bank should provide the user with a relatively uncomplicated way of finding out more about these other concepts, such as hypertextual links.

In fact, this is already being done in several research projects where the goal is to create term banks to be used by computer systems [cf. Meyer/Bowker/Eck 1992a/b, Pearson 19921]. Computer systems are quite similar to non-expert users, in that they need to be told as much as possible about a concept in order to "understanding" it. Some researchers now refer to all the information gathered for a given concept as the definition. Pearson [1992:19], for example, describes how "definitions" are created primarily for use within a machine translation system; these definitions are a series of conceptual characteristics which a computer system can access. Interestingly enough, however, natural language definitions are also included, no doubt to provide human users with a quick overview of the concept.

4.4 CONCLUSION

This chapter has discussed the ways in which the terminologist can view the ITD: as a product, a process, and a tool for users. As a product, the ITD is composed of two main components: the genus and the differentia. While the process of isolating the genus is relatively straightforward, deciding which characteristics to include is more complicated. The terminologist needs to be aware that the characteristics should meet the necessary and sufficient criteria, and that there are tests available (such as Boutayeb's [1990] questions based on "tout le défini et rien que le défini") which can help her confirm the selection of these characteristics.

From the user standpoint, the ITD may be regarded as a multi-purpose tool: along with educating the user, the ITD situates the concept within the concept system, establishes a link between the concept and its term, establishes interlingual equivalences, and contributes to standardization. While not all users need the ITD for all these reasons, the ITD obviously plays a very important role in providing the user with information about the concept.

ITDs are created from the information gathered during concept analysis; therefore, the terminologist needs to ensure that a knowledge-based approach to constructing ITDs takes advantage of the available technology (in this case, CODE). The general technology described in chapter 2 of this thesis has been specifically enhanced to aid in definition construction. This enhancement, to which I contributed as part of my thesis work, is discussed in the next chapter.

Chapter 5	THE TECHNOLOGY: THE CHARACTERISTIC COMPARISON MATRIX

5.0 INTRODUCTION

The introduction to part 2 of this thesis described CODE, the generic knowledge management tool used to help carry out the definition-oriented approach to concept analysis. CODE's general features were described in chapter 2, while its role in the general approach was discussed in chapter 3.

Since definition construction and concept analysis overlap, certain CODE features used during concept analysis are also employed for definition construction. With definition construction in mind, the terminologist can use these features to focus on the two main components of the ITD during concept analysis. As explained in chapter 4, the two main components of the ITD are (1) the generic concept of the concept being defined (denoted by the genus) and (2) the n+s characteristics of the concept being defined (denoted by the differentia). The CODE features which can help the terminologist determine these two components are

1. textual and graphic browsers, which can *automatically display generic-specific relations* can be used by the terminologist to quickly pinpoint the generic concept[54] of the concept being defined, and ensure that definitions of co-concepts all have the same genus (see chapter 2, section 2.2.1);
2. *facets*, which can be used to annotate characteristics, allowing the terminologist to record important information about characteristic names and values (see chapter 3, section 3.1.5);
3. *inheritance mechanisms*, which help the terminologist maintain control over the quality of the characteristic content and syntax, and avoid unnecessary work (see chapter 2, section 2.2.2);
4. the *Characteristic Comparison Matrix (CCM)* can be used both in collecting characteristics and in deciding which ones should be considered n+s and thus included in the ITD.

[54] As was noted in chapter 4 (section 4.2.1), it is possible for a concept to have more than one generic concept. This phenomenon lies outside the scope of this thesis, however, and it is assumed that each concept to be defined has only one generic concept.

Since the first three features have already been discussed in detail, this chapter describes only the Characteristic Comparison Matrix (CCM). Its evolution (from a simple feature in CODE2 to a powerful feature in CODE4), its functionality, and its use in facilitating definition construction are all documented here.

5.1 THE CHARACTERISTIC COMPARISON MATRIX (CCM)

The Characteristic Comparison Matrix was created to satisfy the need for a more effective means of comparing characteristics than was available in CODE2. It has evolved considerably in its relatively short lifespan.

5.1.1 The Evolution of the CCM

The CCM was not originally a feature of CODE2, which was in development from 1988 to August 1991. Initially, it was thought that terminologists using CODE2 would be able to compare the characteristics of co-concepts during an eventual definition construction phase[55] using separate on-screen windows called Concept Descriptor (CD) Views. The general screen layout as the terminologist would have seen it is shown in Figure 30.

Each CODE2 CD View showed all the recorded information for one concept:

— background characteristics (top panes)

— space for a definition

— conceptual characteristics (under the heading Conceptual Information)

— linguistic characteristics (under the heading Linguistic Information)

This format would make it necessary for the terminologist to "filter out" irrelevant information in order

[55] Because of constraints of time and resources, definition construction never became a part of the activities of COGNITERM researchers. This thesis thus represents the first attempt to construct definitions using the CODE system.

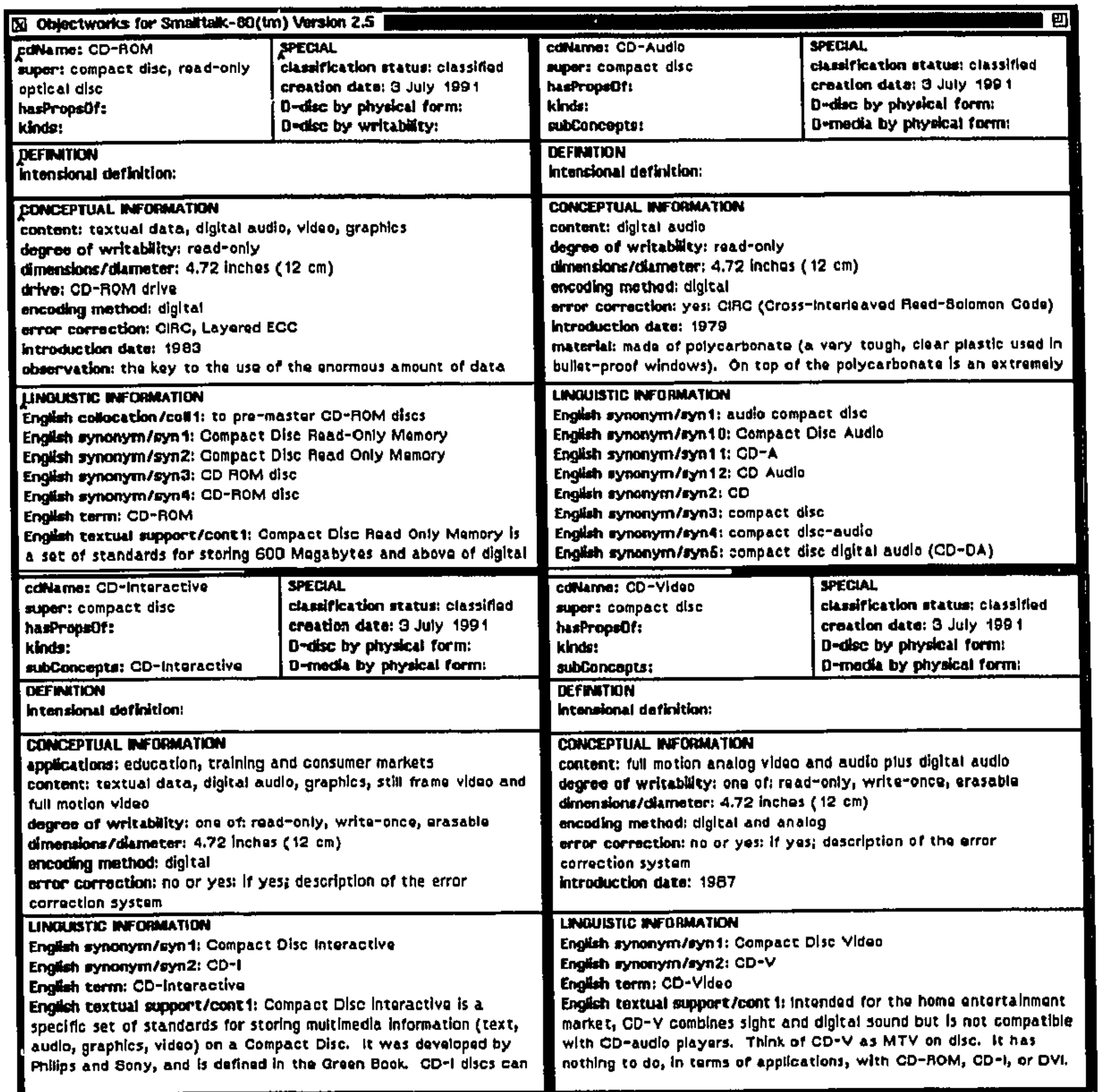


Figure 30. CODE2 CD Views for the co-concepts **CD-ROM, CD-Audio, CD-Interactive,** and **CD-Video,** displayed on-screen for characteristic comparison.

to concentrate on comparing conceptual characteristics.

The CCM was devised in the summer of 1991 as the result of a suggestion from Dr. Geneviève Mareschal, a professor at the School of Translation and Interpretation at the University of Ottawa. While viewing a demonstration of CODE2, Dr. Mareschal commented that, with all the characteristics in such a well-structured format, it would be that much easier to compare them if they were displayed in a tabular interface, similar to that of a spreadsheet, with only the conceptual characteristics of co-concepts visible. And thus began the Characteristic Comparison Matrix (CCM). (See Appendix A for the specifications provided to the programmer[56] who wrote the CCM code.)

Compared to the CODE4 CCM, the CODE2 CCM is rather primitive; even so, it was remarkably useful. It required 10 hours' programming time (very little by programming standards) and immediately underwent testing by me to see what improvements could be made, as I was becoming interested in definitions. The CODE2 CCM was a prototype whose main function was to enable the COGNITERM researchers to analyze it and, from this analysis, suggest improvements for the CODE4 version.

5.1.2 The CODE2 CCM

The CODE2 CCM was the last feature to be developed for that version of CODE. Implemented in August 1991, its primary function was to aid the terminologist in comparing characteristics, both during concept analysis, to ensure that the conceptual information was consistent and accurate in both content and syntax, and after concept analysis, to help with decisions on which characteristics to include in ITDs.

5.1.2.1 How the CODE2 CCM Worked

The CODE2 CCM could be opened for any concept selected in CODE2 to immediately display the differentiating[57] characteristics of the concept and its co-concepts. The CCM (Figure 31) would appear in the form of a table, looking somewhat like a spreadsheet. The selected concept (in

[56] The programmer was Ken Iisaka, at the time an undergraduate student at the University of Ottawa's Department of Computer Science. Mr. Iisaka also wrote the code for the CODE4 CCM.

[57] Differentiating characteristics are those of which have different values for at least two of a set of co-concepts.

Objectworks for Smalltalk-80(tm) Version 2.5

Supers: COMPACT DISC	CD-AUDIO	CD-ROM	CD-Interactive	CD-Video	Compact Video Disc	OROM	DataROM	erasable disc	WORM disc
content	digital audio	textual data, digital audio, video, graphics	textual data, digital audio, graphics, still	full motion analog video and audio plus	textual data and full motion analog video	one or more of: textual data, audio,	one or more of: textual data, audio,	one or more of: textual data, audio,	one or more of: textual data, audio,
degree of writability	read-only	read-only	one of: read-only, write-once,	one of: read-only, write-once,	one of: read-only, write-once,	one of: read-only, write-once,	one of: read-only, write-once,	erasable	write-once
dimensions/diameter	4.72 inches (12 cm)	4.72 inches (12 cm)	4.72 inches (12 cm)	4.72 inches (12 cm)	4.72 inches (12 cm)	5.26 inches (13 cm)	5.26 inches (13 cm)	5.25 inches (13 cm)	3.6, 4.7, 6.26, 8, 12, or 14 inches (most
encoding method	digital	digital	digital	digital and analog	digital and analog	digital or analog or both	digital or analog or both	digital or analog or both	digital or analog or both
error correction	yes: CIRC (Cross-Interleaved	CIRC, Layered ECC	no or yes: if yes; description of	no or yes: if yes; description of	no or yes: if yes; description of	no or yes: if yes; description of	no or yes: if yes; description of	no or yes: if yes; description of	no or yes: if yes; description of
introduction date	1979	1983	1986	1987	n/a	n/a	n/a	n/a	n/a
material	made of polycarbonate (a very tough,	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
standard	Red Book	Yellow Book, High Sierra Group	Green Book	one of Red Book, Yellow Book, Green	one of Red Book, Yellow Book, Green	one of Red Book, Yellow Book, Green	one of Red Book, Yellow Book, Green	one of Red Book, Yellow Book, Green	none
storage capacity	up to 74 minutes of stereo sound	total: 600 MB, user data: 550 MB	120 MB	60 minutes of sound per side, 60 minutes of	At a speed of 900-1,800 rpm, it is			1 gigabyte and more	200 MB - 1.2 GB
drive	n/a	CD-ROM drive	n/a	n/a	n/a	n/a	n/a	n/a	n/a
observation	n/a	the key to the use of the enormous	n/a	n/a	n/a	n/a	n/a	n/a	n/a
applications	n/a	n/a	education, training and	n/a	n/a	n/a	n/a	n/a	n/a

Figure 31. CODE2 CCM for the concept **CD-Audio.**

this case **CD-Audio**) was found in uppercase letters at the top of the second column from the left. The generic concepts of the selected concept were listed in the top-left cell labelled "Supers"[58] (the plural is used because, as mentioned previously, a selected concept may have more than one generic concept). Since **CD-Audio** had only one generic concept, i.e. **compact disc**, only **compact disc** was listed in the Supers cell.[59]

The co-concepts of **CD-Audio**, listed along the top row of the CCM, were **CD-ROM, CD-Interactive, CD-Video, Compact Video Disc, OROM, DataROM**, and **erasable disc**. The differentiating (in fact, the characteristic names) of these co-concepts were listed along the left side of the CODE2 CCM, including ***content***, ***degree of writability***, ***dimensions/diameter***,[60] and ***encoding method***. Those characteristics with values identical for all the co-concepts would not appear in the CCM, as only differentiating characteristics were displayed. For example, because all the concepts in Figure 31 had the identical characteristic ***physical form:disc*** (inherited from **optical disc**, the generic concept of **compact disc**), it did not appear in the CCM.

The remaining cells in the CCM, to the right of the characteristic name cells and below the concept cells, were "value cells." Each value cell of the CODE2 CCM was either filled with a value, left blank, or marked "n/a" or "BLOCKED".

— If the characteristic had been added to or inherited by a concept, and had a value, that value was found in the appropriate cell. For example, in Figure 31, **CD-Audio** has the characteristic ***degree of writability:read-only***: the characteristic name (***degree of writability***)

[58] In keeping with AI terminology, generic concepts are called superconcepts in CODE.

[59] If the selected concept had more than one generic concept, the user was required to select a generic concept. This generic concept would dictate which co-concepts would appear in the CCM. It would appear in the CCM in upper case letters, while any others would be in lower case letters.

[60] Because it was not possible to place characteristics in hierarchies in CODE2, characteristics which were felt to belong to a hierarchy had to be named along with the generic characteristic to demonstrate the fact. For example, in the optical storage TKB, ***diameter*** and ***thickness*** were two characteristics that were considered to be kinds of ***dimension***. Thus in CODE2 they had to be named ***dimensions/diameter*** and ***dimensions/thickness***.

is in the left-hand column, while the value (*read-only*) is in the value cell under **CD-Audio.** The cells were scrollable, so that if the whole value did not fit in the cell frame, the user could easily view the rest of it.

— If the characteristic name had been entered for a concept, but had no value at the time, the value cell would be blank. For example, since no value for ***storage capacity*** had been recorded for either **OROM** or **DataROM,** the cell was blank.

— If a given concept had a characteristic added which the others did not have, the value cells of the others would indicate "n/a". For example, the characteristic ***applications*** had been added for **CD-Interactive** only; the value was recorded for **CD-Interactive,** while the value cells for the remaining co-concepts were marked "n/a."

— If the characteristic had been blocked for a given concept, the value cell displayed the word BLOCKED. For example, although the concept **ostrich** would inherit the characteristic ***method of transportation:flying*** from its generic concept **bird,** this would have been blocked (since ostriches cannot fly), and would appear as such in the CCM.

Once the prototype CCM had been developed, I began a thorough analysis of it. The resulting suggestions as to how the CCM could be improved for CODE4 [cf. Eck 1992] are detailed in the next section.

5.1.2.2 Proposed Improvements for the CODE4 CCM

The CODE2 CCM was quite a remarkable feature, thought to be the only one of its kind in terminology or AI research.[61] Although the CODE2 CCM was helpful in presenting the terminologist with a quick, concise view of the conceptual characteristics, it soon became obvious that

[61] Timothy Lethbridge, Ingrid Meyer [personal communication].

added features would make it even more useful. The additions, namely editability, the ability to display all characteristics, the ability to distinguish necessary and typical characteristics, and masking capabilities, are discussed below.

Editability

The CODE2 CCM showed great promise as an interface for knowledge acquisition. One major problem, however, was that it was not editable.[62] In CODE2, the characteristic values could only be modified in the Characteristic Browser. What was worse, these modifications did not appear immediately in the CCM; rather, the CCM had to be closed and then reopened before it could display the modified information.

This restriction limited the terminologist to using the CCM as a *passive tool*; it could only be used to verify information, not to modify it. The more the CODE2 CCM was used, the more this limitation became apparent: since it was in the CCM that the terminologist could best see the characteristic names and values and ensure their consistency, there was a natural desire to use it to make any necessary modifications. Thus, I proposed that (a) the CCM should be made editable, and (b) any changes made to the information in the CCM should be immediately reflected in the Characteristic Browser and graph, and vice-versa.

Ability to display all characteristics

One option which was proposed for the CCM in the "CCM Specifications" (see Appendix A), but not implemented in the prototype, was a facility to allow the user to view identical characteristics

[62] For a feature to be considered editable, the user must be able to modify the concepts or characteristics viewed in the feature and have these modifications reflected in the rest of the knowledge base. For example, the Characteristic Browser is editable: modifications made to its concepts or characteristics are reflected in the graph and CCM and any other browser where that information can be accessed.

as well as differentiating ones: a window entitled "Characteristics Identical in Co-concepts" would appear and list the characteristics and values in question. Figure 32 shows what might appear in such a window for **CD-Audio** and its co-concepts (i.e. those seen in Figure 31).

Characteristics Identical in Co-concepts

- **physical form: disc**
- **recording technology: optical**

Figure 32. Subwindow showing the identical characteristics of co-concepts (i.e. characteristics not displayed in the CCM).

It soon became clear that such a window could prove to be quite useful. It is likely that, once concept analysis has been completed, the terminologist will not feel the need to view characteristics in the CCM that are identical for all concepts. However, the terminologist also needed to be able to use it *during* definition-oriented concept analysis. As such, the CCM should be able to display *all* the available conceptual information at any given time. Viewing all the characteristics of co-concepts can be an especially valuable option if the terminologist is using a "top-down" approach (i.e. working from more general concepts to more specific ones) during concept analysis: a characteristic (name and value) may initially inherit from a generic concept to all its specific concepts, giving the impression that it is identical for those co-concepts. However, the characteristic value may not actually be identical for all concepts at that level (for example, if the information has not been finalized). The problem with a CCM that shows only differentiating characteristics is that a terminologist examining such a CCM during concept analysis may not be aware that there are characteristics which require further research. The result may be that incomplete or incorrect information is used to construct ITDs.

For example, the co-concepts **CD-ROM XA** and **DVI** initially inherit the characteristic ***content:one of:textual data, audio*** from their generic concept **CD-ROM**. Assuming that concept analysis is in progress and that the terminologist has yet to finalize characteristic values, a CCM displayed at this point for these two concepts would not show this characteristic. If this characteristic does not appear on the CCM, the terminologist may neglect to look further for values in the documentation, or may neglect to question the expert about it. There are two possible endings to such a scenario: at best, the terminologist eventually discovers that more work needs to be done when concept analysis was thought to have been complete; at worst, the characteristic values remain identical, and inaccurate information may be used in ITD construction.

Ability to distinguish necessary characteristics from typical ones

Since only necessary characteristics are considered candidates for n+s characteristics, and thus for inclusion in constructed ITDs, it would be helpful to be able to distinguish necessary characteristics from typical ones in the CCM.

1) Distinguishing characteristics using the STATUS facet: What was needed was a means of distinguishing necessary characteristics from typical ones so that each group could be masked out separately. The most effective way to accomplish this in CODE2 appeared to be via the *STATUS* facet. While CODE2 offered the user several default values for the *STATUS* facet (*SUGGESTED, PROPOSED, IN PROGRESS, UNDER REVIEW*, and *ACCEPTED*), the user could also add different values. A simple solution would be to give necessary characteristics the status of *N*, and typical characteristics that of *T*. Figure 33 shows how one of **CD-Audio**'s characteristics, ***content:digital audio***, could have been given the status of *N* in CODE2.

cdName: CD-Audio
super: compact disc
hasPropsOf:
kinds:
subConcepts:
inheritPropsTo:

PropDesc: type of information that can be stored on medium, i.e. textual data, digital audio, graphics, still pictures, video
References: SAFFADY88:6
Status: N
Category: Conceptual Information
Name: content

digital audio

Figure 33. Status facet displayed for **CD-Audio**'s characteristic ***content:****digital audio* (within the CODE2 Characteristic Browser).

2) Distinguishing characteristics directly on the CCM: Since it seemed quite likely that the

terminologist would consult the CCM often when deciding on the status of a given concept's characteristics, it would also be useful to have some way of indicating the characteristic's status from the CCM itself. One possible procedure was to (a) decide on each characteristic's status (i.e. *N* or *T*) while examining the CCM; (b) somehow put an *N* or *T* label on each characteristic; and (c) close and then re-open the CCM with the typical characteristics masked out. This would allow the terminologist to focus her energies on the necessary characteristics when it came time to decide which ones to include in the ITD. The final CCM, with masked-out typical characteristics, would look like the one shown in Figure 34.

Masking capabilities

The suggestions above indicate that a mask for the CCM would be extremely helpful. For example, the terminologist may have wished to "hide" all the characteristics deemed typical (i.e. with the facet *STATUS:T*). The terminologist may also have wanted to remove any concepts interfering with the comparison of co-concepts during definition construction, such as the still-evolving characteristics of **Compact Video Disc**, **DataROM**, and **OROM** (as can be seen in Figure 34, most of their characteristics have been inherited). Unfortunately, a CCM such as the one seen in Figure 34 was not possible because no mask feature existed for the CODE2 CCM. Thus, the terminologist had no way of temporarily masking out characteristics or concepts either individually or as a group.

Supers: COMPACT DISC read–only optical	CD–ROM	CD–Audio	CD–Interactive	CD–Video	Compact Video Disc	OROM	DataROM	erasable disc
content (n)	text and digital audio	digital audio	combination of digital audio, computer–	full motion analog video and audio plus digital audio	full motion analog video and computer–	one or more of: textual data, audio, graphics, still	one or more of: textual data, audio, graphics, still	textual data and graphics
degree of writability (n)	read–only	read–only	one of: read–only, write–once, erasable	one of: read–only, write–once, erasable	one of: read–only, write–once, erasable	one of: read–only, write–once, erasable	one of: read–only, write–once, erasable	erasable
encoding method (n)	digital	digital	digital	digital and analog	digital and analog or both	digital and analog or both	digital and analog or both	digital and analog or both
standard (n)	Yellow Book, High Sierra Group Standard	Red Book	Green Book	one of Red Book, Green Book, Yellow Book	one of Red Book, Green Book, Yellow Book	one of Red Book, Green Book, Yellow Book	one of Red Book, Green Book, Yellow Book	one of Red Book, Green Book, Yellow Book
storage capacity (n)	total: 600 MB, user data: 550 MB	up to 74 minutes of stereo sound	120 MB	60 minutes of sound/side, 60 minutes of	at a speed of 900–1800 rpm, it is capable of running			1 gigabyte or more

Figure 34. CODE2 CCM with typical characteristics masked out and only necessary characteristics showing (suggested view).

5.1.3 The Initial Development of the CODE4 CCM

As was noted earlier, little time was spent developing the CODE2 prototype CCM. CODE4 was in progress, and the programmers were concentrating on developing and testing it. It was more productive to invest time and effort on the new CCM; the time spent on the prototype CCM was in the form of careful testing and analysis.

After the CCM modifications were proposed, a period of intensive discussion between the COGNITERM team (including me) and the CODE programming team took place. It was agreed that the CODE programmers would develop an initial version of the CODE4 CCM incorporating my suggestions and their own ideas for improvements. This CCM would then be reviewed by me to see what changes were needed. This preliminary programming and analysis took place in the fall of 1992, and allowed me to play an active role in the development of the final CCM. There were times when the CCM programmer (Ken Iisaka) and I would sit computer-by-computer, testing and analyzing the CCM: I would suggest a modification, and get the heart-warming response of "Give me ten minutes...." Ten minutes later, my wish had come true, and the modification had been implemented. Timothy Lethbridge, the head programmer for CODE4, also played a crucial role in ensuring that the CCM interface corresponded to that of the other CODE4 features; he, Ken and I met several times to ensure that the needs of both the COGNITERM team and the CODE team were being met. It was an exciting time in the development of CODE4, and the final result—a true team effort—was a carefully conceived CCM.

5.1.4 The CODE4 CCM

Although it has many new features, the CODE4 CCM works on the same basic principle as the earlier version, allowing the terminologist to view and compare the characteristics of concepts in tabular form. However, many of the modifications suggested in section 5.1.2.2, along with others provided by the CODE4 designers, have been incorporated into this version, making it much more powerful and "terminologist-friendly," affording the terminologist a great deal of control over the recorded information.

5.1.4.1 Using the CCM for Differentiating Characteristics

Once a concept has been selected in a browser (either textual or graphical), the CCM command is given from the CODE menu. The browser in which the concept is selected is the driving browser, while the CCM is the driven browser.[63] In keeping with the CODE4 "interface philosophy," the CCM can be consulted either (a) detached from its driving browser (i.e. in a separate window) or (b) attached directly to its driving browser (below or to the right). Figure 35 shows a detached CCM for the concept **CD-Audio** as it first appears when generated.

Unlike the CODE2 CCM, the CODE4 CCM does not immediately show the differentiating characteristics for the selected concept and its co-concepts. Instead, it initially shows only the selected concept[64] and its characteristics, along with the concept's generic concepts[65] and dimensions (if any). The various elements are shown in Figure 35:

— along the top: generic concept, e.g. **compact disc** (s). The 's' indicates the normal generic-specific relation between the selected concept and its generic concept.

— below the generic concept: the selected concept. Note that the selected concept is also reflected in the CCM label (top left of window)

— far-left column: names of the selected concept's characteristics

[63] Browsers are explained in more detail in chapter 2, section 2.2.1.

[64] It is possible to select more than one concept for display in the CCM; however, for simplicity, it is assumed throughout this chapter that only one concept is selected unless otherwise stated.

[65] Remember that a concept may have more than one generic concept.

Matrix of statements of CD-Audio and coordinates

	> compact disc: (s)
	CD-Audio
physical standard	Red Book
logical standard	Red Book
sector layout	CLV
error correction	CIRC
diameter	4.72 inches (12 cm)
thickness	1.2 mm
writing principle	ablation
content	audio
available recording surfaces	one
storage capacity	up to 74 minutes of audio
physical form	disc
degree of writability	read-only
physical appearance	polycarbonate (a tough, clear plastic) covered with a thin layer of aluminum and a layer of clear resin
data signal type	digital
method of reading	light reflection
recording technology	optical

Figure 35. CCM for the concept **CD-Audio**, as seen when first generated.

— remaining cells: values of the selected concept's characteristics

Based on this initial information, the terminologist can further specify what else should appear in the CCM. In other words, the CCM's content is very user-specifiable, as proposed in section 5.1.2.2. For example, if the terminologist wishes to see the differentiating values of the selected concept and its co-concepts, she simply clicks on the appropriate generic concept. Figure 35 shows that **CD-Audio** has only one generic concept (**compact disc**). When the terminologist clicks on **compact disc**, the CCM shows the differentiating characteristics of **CD-Audio** and its co-concepts. The resulting CCM is shown in Figure 36.

Matrix of statements of CD-Audio and coordinates

	compact disc: (s)						
	CD-Audio	CD-ROM	CD-Interactiv	CD-Video	OROM	DataROM	Compact Vide
physical standard	Red Book	Yellow Book	Green Book	Blue Book	one of: Red Book, Yellow Book, Green	one of: Red Book, Yellow Book, Green	one of: Red Book, Yellow Book, Green
logical standard	Red Book	Group Standard, ISO 9660	Green Book	Blue Book	one of : Red Book , Yellow Book , Green	one of : Red Book , Yellow Book , Green	one of : Red Book , Yellow Book , Green
sector layout	CLV	CLV	CLV	CLV	CAV	CLV	CLV
error correction	CIRC	CIRC, Layered ECC	yes or no	yes or no	yes or no	yes or no	yes or no
diameter	4.72 inches (12 cm)	4.72 inches (12 cm)	4.72 inches (12 cm)	4.72 inches (12 cm)	5.25 inches (13 cm)	4.72 inches (12 cm)	4.72 inches (12 cm)
content	audio	textual data, audio, graphics, still	one or more of: textual data, audio,	full motion analog video and audio plus	one or more of: text, audio,	one or more of: text, audio,	textual data and full motion analog
storage capacity	up to 74 minutes of audio	total: 600 MB, user data: 550 MB	120 MB	60 minutes of sound per side, 60	•	•	At a speed of 900–1,800 rpm, it is
physical appearance	polycarbonate (a tough, clear plastic)	•	•	•	•	•	•
data signal type	digital	digital	digital	digital and analog	digital	digital or analog or both	digital and analog
method of reading	light reflection	•	•	•	•	•	•

Figure 36. CCM for **CD-Audio** and its co-concepts under **compact disc**, showing differentiating characteristics.

5.1.4.2 Additional CODE4 CCM Features

While the CCM shown in Figure 36 facilitates the viewing of concepts and their characteristics, the terminologist is not obliged to work with the information as it is presented. The new features—editability, ability to display characteristics and concepts in a variety of modes, masking capabilities, and rearranging facilities—allow the terminologist to manipulate the information to her advantage.

Editability

The CODE4 CCM is editable, allowing the terminologist to modify characteristic values at any time. As was proposed in 5.1.2.2, any modification made to a characteristic value in the CCM now automatically appears in the Conceptual Info browser, and vice versa. Along with knowledge revision, the CCM can now be used for knowledge acquisition.

Ability to display concepts and characteristics in a variety of formats

Up to this point, it has been assumed that the terminologist wishes to see only co-concepts and differentiating characteristics. Fortunately, the CODE4 CCM does not restrict the terminologist to viewing only this combination. The CCM Format is a feature which was derived from the suggestion that the user be able to view all of a concepts' characteristics. The CODE4 programmers took this suggestion one step further: the terminologist can adjust the CCM format so that concepts and characteristics appear in a variety of formats. This may be done before or after the CCM is opened from the CCM Format dialog box (Figure 37).

The CCM Format lets the terminologist control what appears in the CCM. There are three main format parameters, each of which can be set several ways: (1) Subjects (i.e. Concepts); (2) Properties (i.e. Characteristics); and (3) the Property Computation Operator (PCO). These three parameters are all taken into account by the CCM when it computes what should appear in its cells at various stages.

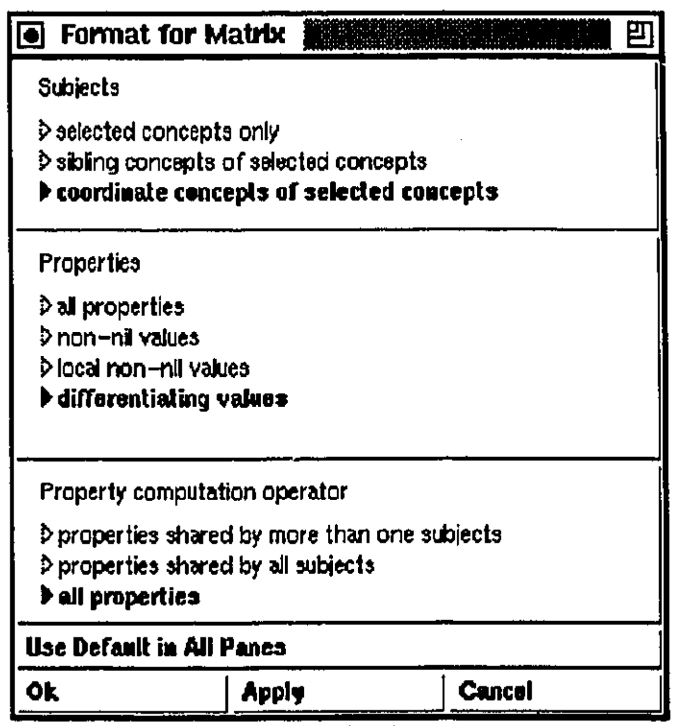


Figure 37. CCM Format dialog box.

Of greatest interest to the terminologist are the Subjects (Concepts) parameter and the Properties (Characteristics) parameter. By specifying different Concepts formats, the terminologist can view

— one or more arbitrarily selected concepts (useful for viewing a concept and its generic concept, for example, as seen in Figure 38, where **CD-ROM** and its generic concept **compact disc** are displayed)

Matrix of statements of (compact disc, CD-ROM)

	compact disc	CD-ROM
physical standard	one of: Red Book, Yellow Book, Green Book, Orange Book; if Yellow Book, also	Yellow Book
logical standard	one of : Red Book , Yellow Book , Green Book , Orange Book ; if Yellow Book ,	{High Sierra Group Standard, ISO 9660 }
sector layout	CLV	CLV
error correction	yes or no	{CIRC, Layered ECC }
diameter	4.72 inches (12 cm)	4.72 inches (12 cm)
thickness	1.2 mm	1.2 mm
rotation speed		
writing principle	ablation	ablation
content	one or more of: text, audio, graphics, still pictures, and motion video	textual data, audio, graphics, still pictures, and motion video
available recording surfaces	one	one
storage capacity		total: 600 MB, user data: 550 MB
physical form	disc	disc
degree of writability	one of: read-only, write-once	read-only
physical appearance		
data signal type	digital or analog or both	digital
laser type		
method of reading		
recording technology	optical	optical

Figure 38. CCM for the arbitrarily selected concepts **compact disc** (generic concept) and **CD-ROM** (specific concept), with all characteristics showing.

— sibling[66] concepts of the selected concept

— co-concepts of the selected concept

By specifying different characteristic formats for the CCM, the terminologist can choose, for example, to view all the characteristics of a concept and its co-concepts, or only the differentiating characteristics. Such a feature allows the terminologist to view all the recorded information at any time during concept analysis in order to verify, for example, that characteristics are consistent in content and syntax, or to see whether further research is required for a particular concept.

Masking capabilities

The CODE4 CCM mask allows the terminologist to "hide" both concepts and characteristics, and thus focus on the remaining information. The terminologist can mask out an individual concept or characteristic by clicking on it; for example, to hide a concept for which little information has been collected. Concepts and characteristics may also be masked out as a group. For example, once the terminologist has decided which characteristics are typical, she can give their *MODALITY*[67] facet a value of *T (TYPICAL)*, and then mask out all the characteristics with the facet *MODALITY:T*.

Figure 39 shows the CCM from Figure 36 with several concepts (**DataROM, OROM, Compact Video Disc**) and characteristics (***sector layout***, ***diameter***, ***physical appearance***, ***method of reading***) masked out.

Ability to rearrange concepts and characteristics

While it is still not possible to visibly distinguish necessary characteristics from typical ones on the CCM, it is possible to rearrange concepts and characteristics, either alphabetically or according

---

[66] In CODE, sibling concepts are distinguished from co-concepts according to the *dimension* phenomenon (discussed briefly in the introduction to this thesis). Since dimensions lie outside the scope of this thesis, I assume here that there is no difference between co-concepts and sibling concepts.

[67] In CODE4, *STATUS* has been renamed *MODALITY*. While this term is unknown in the terminology literature, it is accepted in the knowledge engineering world [Douglas Skuce, personal communication]. It is used here to indicate whether a characteristic is necessary, typical, or n+s.

Matrix of statements of CD-Audio and coordinates

▸ compact disc: (s)

	CD-Audio	CD-Interactive	CD-ROM	CD-Video
physical standard	Red Book	Green Book	Yellow Book	Blue Book
logical standard	Red Book	Green Book	(High Sierra Group Standard, ISO 9660)	Blue Book
error correction	CIRC	yes or no	(CIRC, Layered ECC)	yes or no
content	audio	one or more of: textual data, audio, graphics, still frame video and motion	textual data, audio, graphics, still pictures, and motion video	full motion analog video and audio plus digital audio
storage capacity	up to 74 minutes of audio	120 MB	total: 600 MB, user data: 550 MB	60 minutes of sound per side, 60 minutes of video per side
data signal type	digital	digital	digital	digital and analog

Figure 39. CCM for **CD-Audio** and co-concepts with various concepts and characteristics masked out. Compare with original CCM in Figure 37.

to the terminologist's wishes (i.e. to the left/right, to the bottom/top). This is one of the more important CCM features, as it gives the terminologist the power to arrange the display for her own purposes. For example, the terminologist may not wish to mask out typical characteristics; she may instead place them at the bottom of the CCM, leaving the necessary characteristics at the top where she can concentrate on them.

5.2 CONCLUSION

The CODE4 CCM is an improved version of the CODE2 CCM, with a variety of features which make it more powerful and terminologist-friendly. Table 8 provides a brief summary of the modifications to the CODE2 CCM proposed in section 5.1.2.2 and the CODE4 features that resulted from both these proposals and others made by the CODE programmers.

PROPOSED MODIFICATIONS TO CODE2	CODE4 FEATURES
Editability (ability to make modifications to characteristics which would be reflected in the whole TKB)	Editability: modifications made to characteristics in the CCM are reflected in the other browsers and vice versa; concept additions and deletions made in browsers are reflected in the CCM
Ability to display all characteristics, not just differentiating ones	Ability to display characteristics and concepts in a variety of modes: (a) characteristics: all, differentiating, etc.; (b) concepts: coordinate, sibling, arbitrarily selected
Masking capabilities (ability to hide concepts and characteristics as individuals or in groups)	Masking capabilities (ability to hide concepts and characteristics as individuals or in groups)
Ability to visibly distinguish necessary characteristics from typical ones	Ability to rearrange characteristics and concepts within CCM

Table 8. Comparison of features for the CODE2 CCM and CODE4 CCM.

The CODE4 CCM is a powerful tool which helps the terminologist both during and after concept analysis. During concept analysis, it acts as an knowledge-editing tool, allowing the terminologist to

— see all the currently available information, or only those characteristics which differentiate concepts

— identify where more work is required (e.g. where characteristic values need to be modified or updated) and make any necessary modifications

— present to experts the information gathered so far

Once concept analysis has been completed to some degree, the CCM can act as a knowledge-revision tool, allowing the terminologist to

— confirm that the information is correct in both content and form

— show experts the "final product" for verification

— manipulate the information using the masking and rearranging facilities so that eventually only potential n+s characteristics are left to consider for definition construction

Of course, this technology alone is not enough: along with the CCM and other CODE features, the terminologist needs an approach to help her systematically construct definitions. While the approach discussed in chapter 3 provided guidelines for carrying out general concept analysis using CODE, it did not take definition construction into account *per se*. The following chapter describes the modifications and additions to the general approach that were devised to ensure that the terminologist carrying out definition-oriented concept analysis effectively utilizes the CCM and other features of CODE.

Chapter 6 THE APPROACH: DEFINITION-ORIENTED CONCEPT ANALYSIS

6.0 INTRODUCTION

Once the terminologist understands the ITD (as a process, a product, and a tool) and how CODE's features, particularly the CCM, can help her record the components of the ITD, the remaining element needed for ITD construction is an approach.

The general approach to concept analysis outlined in chapter 3 did not take definition construction into account. However, as noted in chapter 1 (section 1.4), definition construction is one of several activities that may follow concept analysis. Thus it seems logical that if additions and modifications that take definition construction into account were made to a general approach, the result would be a definition-oriented approach to concept analysis.

This chapter proposes such an approach, which derives from three sources. First, it is based on a review of the literature on intensional terminological definitions, conducted to discover which elements of the ITD the terminologist should concentrate on, and summarized in chapter 4. Second, it is based on an in-depth analysis of the available technology (CODE) in light of the results of this review, summarized in chapter 5. Third, and most importantly, this proposed approach is based on introspection supported by my prior experience with the COGNITERM project. Simply put, I studied how we had initially carried out general concept analysis using CODE (for the subfield of optical storage media), then tried to determine what needed to be done differently if ITDs had to be to constructed at the end of the analysis.

The result is a *definition-oriented approach to concept analysis*, intended to help the terminologist carry out concept analysis oriented towards the construction of intensional terminological definitions using CODE.

6.1 PROPOSED MODIFICATIONS AND ADDITIONS TO THE GENERAL APPROACH TO CONCEPT ANALYSIS

The various steps of general concept analysis were described in detail in chapter 3, so they are given only briefly mentioned here. It is the proposed modifications and additions, summarized in Table 1, which are discussed in detail. When no changes to the general approach are proposed, the right column of Table 1 is left blank.

PROPOSED DEFINITION-ORIENTED APPROACH TO CONCEPT ANALYSIS	
General Concept Analysis (cf. chapter 3)	**Proposed Modifications and Additions**
Part I	**Carrying Out Preparatory Work**
(a) decide on subject field and languages (b) select introductory corpus providing overview of subject field (c) create sketch of concept graph using CODE (general "concept clusters") (d) focus on one particular subfield (indicated by "concept cluster") if subject field is large (e) consult experts about (i) preliminary concept graph and (ii) possible documentation (f) select principal documentation and compile a bibliography (input bibliography directly into CODE) (g) read subfield documentation and establish a preliminary concept system for subfield NOTE: some generic-specific relations should be clear at this point; confer with experts to verify, using graph	· annotate clearly situated generic concepts using facets · record and annotate characteristic(s) underlying the subclassification of a concept · annotate concepts and characteristics which experts consider fundamental to the concept system
Part II	**Establishing Characteristics Template**
(i) design characteristics template from primary documentation, using top-level template as a starting point (j) confer with experts regarding template; correct omissions or delete characteristics that don't belong	· refer to annotations (facets) for guidance · use CCM to show all characteristics to experts; insert expert comments in appropriate facet

Part III	**Carrying Out Primary Knowledge Acquisition & Revision**
(k) scan principal documentation and extract conceptual and linguistic information (l) revise recorded information; have it verified by experts	· consult CCM frequently, using both "differentiating" and "all characteristics" formats, to monitor information · be sensitive to linguistic, conceptual, and structural clues in documentation · annotate characteristics using context, expert, and modality facets · finalize modality of characteristics with help of experts
Part IV	**Constructing the Draft Definition**
	(m) mask out typical characteristics on CCM using mask (based on ***modality***) (n) examine differentiating characteristics of co-concepts on CCM and select potential n+s characteristics, with help of experts, if possible (o) mask out non-n+s characteristics from CCM, leaving draft definition (genus on graph, differentia in CCM)
Part V	**Testing the Draft Definition**
	(p) test the draft definition using two "tout le défini" questions to see which are n+s, with expert if possible (q) retrieve erroneously masked-out n+s characteristics and mask out those which are not n+s; remaining characteristics will be included in the final ITD (r) write natural language ITD, making any necessary modifications for style; e.g. for **CD-Audio:** *ITD:a read-only compact disc that contains 74 minutes of digital audio.*
Postscript	**Updating**
(t) updating	· if an n+s characteristic is updated, or if a new concept is added which may necessitate modifying characteristics, existing definitions should be checked to see if they need revision (use subgraph and CCM to compare new draft ITD with existing ITD)

Table 9. Proposed definition-oriented approach to concept analysis: the general approach with proposed modifications and additions.

6.1.1 Carrying Out Preparatory Work (Part I)

As Table 1 shows, the preparatory work of definition-oriented concept analysis remains largely the same as for general concept analysis. However, even in these early stages, the terminologist should take care to annotate any information that is recorded about concepts and characteristics. The annotations (explained in more detail below) will eventually help the terminologist decide on the genus and differentia (the two components of the ITD) of each concept to be defined.

What should be annotated?

Concepts that seem to have straightforward generic-specific relations should be annotated as such, since generic concepts will be included in the ITD later. Any characteristic which seems to dictate the division of the selected subfield or play an important role in distinguishing concepts should also be recorded and annotated, since it may well constitute an n+s characteristic which should be included in the ITD.

For example, when doing our initial COGNITERM research on optical storage media, we based our preliminary subfield graph on a hierarchical figure found in Chen [1989:5] titled "Types of Optical Media," which divided the subfield into three main types: read-only, write-once, and erasable media.[68] The figure, which displayed approximately 20 concepts, provided us with numerous generic-specific relations, along with a characteristic which was used to subclassify the concept of **optical storage media**. (Since Chen 1989 only provided the values, i.e. *read-only*, *write-once*, and *erasable*, we named the characteristic ***degree of writability***). In retrospect, each specific concept of **optical storage media** could have been annotated to show (a) that it had a fairly straightforward generic-specific relation and (b) where it had been found (in a hierarchical figure). As well, the characteristic ***degree of writability**:one of:read-only, write-once, erasable* could have been recorded

[68] This seemed to be a very standard division, reflected in other documents.

immediately and annotated to show that it was used in this document (and others) to subclassify a top–level concept, and thus should be considered a necessary characteristic for that concept's specific concepts.

How should concepts and characteristics be annotated?

Facets, which are used to record information about characteristics, are essential to the annotation process. While some existing facets can be utilized to help the terminologist "keep notes" on the characteristics being recorded, others must be added:

(a) for concepts: the background characteristic[69] ***superconcept*** should be used to keep track of a given concept's generic concept.[70] In CODE, the generic concept appears as the ***superconcept*** value. For example, when **CD–Audio** is added as a specific concept of **compact disc**, it automatically has the background characteristic ***superconcept:compact disc***.

Since every piece of recorded information is given a reference in the COGNITERM methodology, the existing *REFERENCE* facet can be used to record where the generic–specific relation was documented (e.g. *REFERENCE:CHEN89:5*). This reference can be made more precise if a *CONTEXT* facet, in which the terminologist notes exactly how the relation between the given concept and the generic concept was indicated (e.g. *CONTEXT:HIERARCHICAL FIGURE OF OPTICAL STORAGE MEDIA SUBFIELD*) is added. This way, the terminologist can explain her selection of the generic concept in case the experts consulted later question the recorded information.

(b) for characteristics: since the characteristic name and value are separate entities, they each need a separate set of facets. Along with a *REFERENCE* facet, they should each have a *CONTEXT* facet

[69] Background characteristics are discussed in chapter 3 (section 1.4).

[70] If the terminologist does not know a given concept's generic concept, the concept can be placed under a "dummy" umbrella concept until the concept's exact place in the concept system is discovered.

to record the actual context (e.g. list or table of characteristics, defining context[71]) in which the name or value was found. Again, the terminologist can then justify why the information was recorded.

At this stage, the terminologist can concentrate on the characteristic name, since the majority of values will be recorded later during scanning (see section 6.1.3). If values are found for characteristics of top-level concepts, however, they could be recorded. If the terminologist has to invent a characteristic name (as with the example of ***degree of writability*** above), the terminologist's name can be put as a *REFERENCE*, and an explanatory note put in the *CONTEXT* facet.

The existing *MODALITY* facet can be used to indicate whether the whole characteristic (name and value) is *NECESSARY* or simply *TYPICAL*, depending on the clues provided in the documentation about the characteristic. Such clues include characteristics appearing frequently in the documentation, characteristics found in defining contexts, tables, and lists.

Once the terminologist has created a preliminary concept graph and recorded the more obvious characteristics, she should consult the experts[72] and annotate any concepts or characteristics they feel are particularly important. A third facet, *EXPERT*, should be added for both the characteristic name and value; it can be filled in with the experts' names (which should be considered a special type of reference) to show that they have verified the recorded information, or with an added comment, should the need arise. The resulting facets browser can be seen in Figure 40.

At this point, the terminologist has a very preliminary concept system of the subfield, with some annotated concepts and a selection of annotated characteristics. The next step is establishing a top-level template of characteristics.

[71] Defining contexts should also be recorded (as Linguistic Info characteristics) in case they can eventually be used intact or with little modification.

[72] Ideally, the expert should be sitting in front of the system for this and subsequent sessions. If this is not possible, another (although less effective) option is to use printouts of the graphs, characteristics, facets, etc. to present the information to the expert.

6.1.2 Establishing A Characteristics Template (Part II)

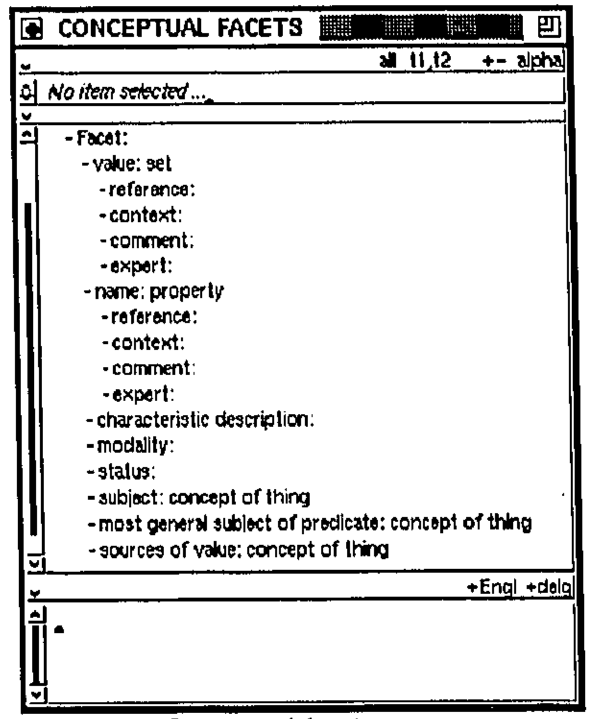


Figure 40. Conceptual facets.

The terminologist can review the characteristics that have been recorded up to this point and decide which ones should be included in the template, using the *CONTEXT*, *MODALITY*, and *EXPERT* facets to help her decide. The characteristics included in this top–level template will be the ones that describe the subfield's top–level concepts.

As the terminologist works her way down the subfield hierarchy, the template may have to be modified to include additional characteristics which serve to divide up subhierarchies or differentiate concepts at those lower levels. The process of deciding on these characteristics remains the same, however: the terminologist must take into account clues in the documentation, along with expert opinion. Expert opinion is valuable at this point. It is often difficult for the terminologist to attain a clear idea from the documentation what the proper dimension of the subfield (the documentation often represents as many points of view as there are documents), and of which characteristics dictate the division. For example, while working on the optical storage TKB, we enlisted the help of Mr. Don Slaunwhite, an optical storage expert from Corel Systems Corporation. He was quickly able to tell us that a particular characteristic that we had been using to divide up the optical storage media subfield (***degree of writability***) was actually not so important, and that we would be better off dividing it up according to ***form*** (as discs, cards, film)

Because of the importance of characteristics, the general approach suggests conferring with subject–field experts again at this point to verify the validity of the characteristics template. Since this template is embellished for the lower-level concepts in the hierarchy, errors or omissions could end up wasting the terminologist's time (in that she might have to go back and redo the template from the top). A faulty template could also result in the terminologist constructing definitions from erroneous or inadequate information. The example of ***degree of writability*** illustrates this point: had we continued researching optical storage media without consulting a subject-field expert, our subfield would have been inappropriately divided up, and the eventually constructed definitions would have been erroneous.

The Characteristic browser can be used at this point to display the template in the form of a list of characteristics for one of the top concepts. The experts may find it easier to consult the Characteristic Comparison Matrix, however, as it displays characteristics in an easy-to-read, well-structured format. The terminologist can exploit the ability of the CCM to show subject-field experts *all* the characteristics identified for any set of concepts or, for example, to show which characteristics have been assigned to a particular set of co-concepts. Once the experts have commented on the top-level template, the terminologist can make any necessary modifications to the characteristics and note the experts' opinions in the *EXPERT* facet.

6.1.3 Carrying Out Primary Knowledge Acquisition & Revision (Part III)

The next step is the scanning of documentation, during which the terminologist records any new concepts and extracts information about the concepts and their terms from the subject-field documentation. This information is recorded in the TKB in the form of linguistic and conceptual characteristics, as explained in chapter 3 (sections 3.1.2 and 3.1.3).

The CCM can be used at this stage for knowledge acquisition: since the CCM is editable, the terminologist may use it rather than the Characteristic browser to directly add or modify characteristic values. Since the terminologist is keeping definition construction in mind, in which she will compare the characteristics of co-concepts, she will likely want to gather information for a set of co-concepts (i.e. treating one set of co-concepts at a time). The CCM, with its ability to display co-concepts and their characteristics, is much better suited for collecting information in this manner than the Characteristic browser. As well, the CCM characteristic format can be set to either "differentiating mode" or "all mode," allowing the terminologist to monitor the progress of her knowledge acquisition and to ensure that the content and syntax of the characteristic values are consistent.

The terminologist should remain alert for clues in the documentation, including *linguistic* clues, such as defining contexts, repeated phrases such as "the most important feature of X is Y" or "the main difference between X and Y is..."; and *structural* clues, e.g. tables, lists, and concept system diagrams. These clues may provide direct pointers to n+s characteristics. As stated above, the characteristic values should be annotated whenever they are modified using the ***REFERENCE***, ***CONTEXT***, and ***MODALITY*** facets.

Along with the Characteristic browser and the CCM, the terminologist should take advantage of CODE's inheritance mechanisms (cf. chapter 2, section 2.2.2 and chapter 5, section 5.0) as much as possible during the scanning process. This allows the terminologist to (a) avoid repeatedly inputting the same information, and (b) ensure, to a certain degree, that the characteristic values of co-concepts are consistent in content and syntax. As noted in chapter 4 (section 4.2.2), consistent characteristics are crucial to the construction of good ITDs.

If the terminologist finds a characteristic which is not included in the initial template but which now appears to be important to describing the concepts, she should add it to the template at the appropriate level (after consulting with experts, if possible), with the appropriate annotations. Typically, more characteristics are added to the template as the analysis proceeds down through the hierarchy.

Once the terminologist feels that scanning is complete, she should review the collected information and try to establish each characteristic's preliminary ***MODALITY*** (*NECESSARY* or *TYPICAL*). Certain rules of thumb can be followed for this process:

— characteristics for which the values have been modified or added for all co-concepts can be considered potential necessary characteristics, since they obviously serve to differentiate the concept from its generic concept and from its coordinates

— characteristics found in defining contexts, lists, or tables, or consistently qualified as "the most important" either in documents or by experts, should be considered potential necessary characteristics; the terminologist can check the ***CONTEXT*** facet for each characteristic value to verify the context in which it was found, along with the ***EXPERT*** facet

— characteristics added to the top-level template during scanning can be considered necessary candidates; if the terminologist felt that the characteristic was important enough to be added to the template, it may well be important enough to be included in the definition

— a characteristic that is clearly the basis for subclassifying any concept can be considered a necessary candidate; e.g. the characteristic on which the subclassification of **optical storage media** was eventually based was *form:one of: disc, card, paper, film*. Thus, this characteristic (or rather, a more specific version of it) would likely be included in a definition of any of **optical storage media**'s specific concepts.

Next, the subject-field experts should be consulted and asked to verify the recorded information and the modality of the characteristics. Using the graph, the CCM, and the Facets browser (to view the annotations), the terminologist and experts can finalize (a) the generic-specific relations between concepts, and (b) the modality of the characteristics (necessary or typical). For example, **CD-ROM** has the characteristic ***container:jewel box***. While **CD-ROMs** typically are packaged in jewel boxes, it is not mandatory that this be done. Thus, its ***MODALITY*** is recorded as

TYPICAL.

The CCM is an effective tool at this point for knowledge revision, i.e. for making any necessary final changes to the form and content of the characteristics.

6.1.4 Constructing the Draft Definition (Part IV)

Once the information has been recorded to the terminologist's satisfaction, the task of deciding which necessary characteristics should be included in the definition can begin. As was explained in chapter 4 (section 4.1.2), the second component of the ITD, the differentia, is the set of necessary and sufficient (n+s) characteristics. These characteristics are individually necessary to a concept and collectively sufficient to describe it and differentiate it from its co-concepts.

To begin the task of deciding upon the n+s characteristics, the terminologist should examine the characteristics of a given set of co-concepts, ensuring that the CCM is in differentiating mode. Next, she masks out the typical characteristics, using the facet *MODALITY:T* as the masking criterion. The remaining characteristics displayed on the CCM are those which have been deemed necessary. The terminologist now has to decide which set of these is sufficient to define the concept. At this point, the terminologist must rely on a combination of subject-field knowledge and creativity: by carefully comparing the necessary characteristics for the various co-concepts, she can establish a tentative set of potential n+s characteristics for each co-concept. Ideally, these potential n+s characteristics are then verified by the subject-field experts, as their selection requires judgement based on expert knowledge. The terminologist can mask out the characteristics not considered n+s by clicking on them. (Their ***modality*** should not be changed for the moment, since the n+s characteristics must be tested as a set before they can be deemed n+s, and changes may have to be made.)

With both the generic concept and the potential n+s characteristics for the concept-to-be-defined displayed in the CCM, the terminologist now has, for all intents and purposes, draft definitions for the co-concepts onscreen in one CODE browser.[73] Genus and differentia have finally been brought together. All that remains to be done is the testing of each draft definition to see if it adequately defines the concept.

6.1.5 Testing the Draft Definition (Part V)

The draft definitions can be tested using the "tout le défini" questions of Boutayeb [1990, based on Rey-Debove [1966]], discussed in chapter 4 (section 4.2.2.1):

(1) Are all concepts which have the generic concept (B) and the set of n+s characteristics (C) denoted by the term (A)?

(2) Do all terms (A) denote a concept which has the generic concept (B) and the n+s characteristics (C)?

Ideally, the experts should be involved in this stage of the process, for they are the most qualified to answer these questions.[74] Chapter 4 (section 4.2.1) warned that the answers to these questions may turn out negative, in which case the definition is either too narrow or too broad. Constant reference to the CCM should help the terminologist and experts discover where any problems lie, and which characteristics need to be deleted from or added to the set of n+s characteristics. Those to be deleted can be masked out, while those to be added can be retrieved from the mask so that they reappear in the CCM.

The subject-field experts should be consulted one last time to approve the draft definitions. Any changes they suggest may be made directly to the CCM.

The last step in part V is the actual writing of the natural language ITD. Since the ITD

---

[73] If it proves easier for the terminologist to view the n+s characteristics one concept at a time (rather than viewing the n+s characteristics for all the coordinate concepts) in the CCM, she can either (a) mask out the other concepts by clicking on them, or (b) deselect the generic concept in the CCM so that only those characteristics describing the concept selected in the graph (the driving browser) show.

[74] Because this is likely to be the most time-consuming part of the approach, it is possible that a subject-field expert may not be able to (or may not want to) spend as much time on the information as the terminologist might like.

contains important conceptual information, the constructed ITD should be included as a special conceptual characteristic (named ***ITD***) in the Conceptual Info browser, shown in Figure 41. This would be the most logical place for the user to look; if she wants to know more about the concept (or obtain a better understanding of the elements of the definition), she need only consult the conceptual characteristics below the ITD. For the ITD "value," the generic concept is combined with the n+s characteristics, and the whole is edited to form a natural language statement based on the "definiens" model (cf. chapter 4, Figure 26). For **CD-Audio**, for example, the n+s characteristics ***degree of writability:read-only***, ***content:digital audio***, and ***storage capacity:up to 74 minutes*** can be added to the generic concept of **compact disc** so that the final ITD reads

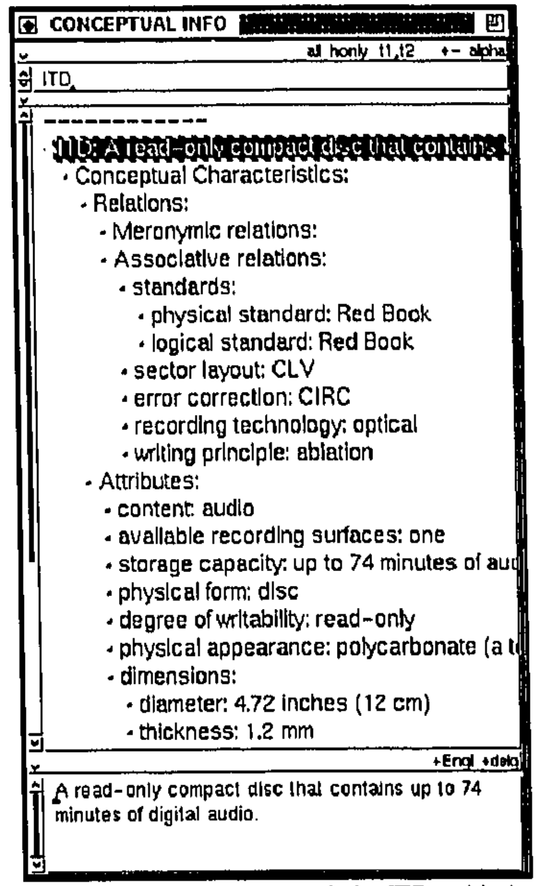


Figure 41. The characteristic ***ITD***, added to the Conceptual Info browser.

ITD:A read-only compact disc that contains up to 74 minutes of digital audio.

Note that the characteristic names may be omitted, and that some characteristics may be combined with the genus to make the definition sound more natural.

6.1.6 Updating Subject-field Information in the TKB (Postscript)

There is one final, ongoing step which the terminologist should keep in mind if she wishes to maintain a "dynamic" TKB. Updating (known as "maintenance" in the software development industry) must be carried out from time to time, especially with constantly evolving subject fields. As new concepts emerge, they should be recorded along with their characteristics. Characteristics for existing concepts may have to be updated at this point as well, since the inclusion of new concepts may necessitate changes to the description of existing concepts. If so, the terminologist should determine whether any new or updated characteristics are also n+s, and ensure that the natural language definitions in which they are found are also updated. The combined graphical browser and CCM can be used once again to either construct a draft ITD for a new concept or compare a revised "draft" ITD of a modified concept with its existing natural language ITD.

6.2 CONCLUSION

While the proposed definition-oriented approach to concept analysis is composed of many individual steps, summarized in Table 1 at the beginning of this chapter, it can be reduced to a small set of important points, discussed below.

6.2.1 Annotating Information

It is very important for the terminologist to carefully annotate both generic concepts and characteristics (name and value) in order to justify later decisions. These annotations include

— the *reference* for each element of information

— the *context* in which the information element is found

— the *modality* of characteristics (necessary vs. typical, then n+s vs. necessary), once this has been determined by the terminologist (relying first on the documentation, then on the experts' knowledge)

— the *experts' opinion* regarding the validity of the information

6.2.2 Developing Sensitivity to Documentation Clues and Expert Opinion

The terminologist must develop a sensitivity to linguistic and structural clues in the documentation that provide information about concept relations and characteristics (names, values, degree of importance), including phrases, section headings in chapters, defining contexts, tables, lists, and figures.

At the same time, the terminologist must develop a sensitivity to what the experts say. In the past, using the general approach, we simply wanted to verify that our characteristic names and values were correct; now we want to determine the most important characteristics. This change in focus may modify how the terminologist communicates with the subject-field experts: it may change the intent of questions, and it may change the way in which the responses are recorded.

6.2.3 Relying on Experts in Matters of Judgement

Following general terminology practice, it is essential to have expert help. This is especially important for definitional matters requiring judgement, such as deciding on the main generic-specific relations, the characteristics responsible for the subclassification of a concept, necessary vs. typical characteristics, and n+s vs. necessary characteristics. While the terminologist may learn a great deal during her research of the subject field, only the extensive knowledge of the subject-field expert can help create the seemingly superficial description of a concept that is the definition.

6.2.4 Using the CCM to Establish a "Definition Mindset"

It is important that the terminologist make the most of the CCM as a knowledge acquisition and editing tool: it gives her considerable control over the information, and, by displaying sets of co-concepts with their characteristics, it helps her establish a kind of "definition mindset" during both concept analysis and expert help sessions. Of course, this does not mean that the CCM should be used to the exclusion of other CODE features such as the graphical and textual browser; they all have their uses and should complement each other.

6.2.5 Understanding the Non-sequential Nature of this Approach

Although this approach is laid out is sequential steps, the terminologist must understand that there is a great deal of overlap in the various activities that occur during definition-oriented concept analysis. Some steps will not be completed until the whole project is done. For example, although some generic-specific relations must be determined near the beginning of the project in order to determine subfields, the terminologist will likely find herself recording other generic-specific relations during the scanning process as well. Characteristics may be chosen for the top-level template before scanning is begun, but others will appear later. The terminologist must be able to adapt to tailor her approach to the flow of information, and perform various steps simultaneously.

Part 2 of this thesis explained how the terminologist should be aware of the ITD as a process, a product, and a tool. It discussed how several of CODE's features, particularly the CCM, can help the terminologist record the information necessary for building ITDs. Finally, it proposed a definition-oriented approach to concept analysis, a modified version of the general approach discussed in chapter 3.

To discover whether this proposed approach yields good ITDs, I needed to test and analyze it in a practical situation, i.e. in a terminology research project in which definition-oriented concept analysis is carried out and ITDs are constructed. A description of such a project, along with an analysis of how the proposed approach fared, are the subject of part 3 of this thesis.

PART 3

TESTING THE APPROACH AND ASSESSING THE RESULTS

Part 3 of this thesis describes a practical test of the definition-oriented approach to concept analysis proposed in chapter 6, and then examines the results.

Chapter 7 details how the test was conducted on 16 concepts in the subject field of **lasers**. It explains what happened as I carried out the approach step-by-step on three different levels of the subject field (**laser**, kinds of **laser**, and kinds of **neodymium laser**).

Chapter 8 provides an assessment of the results of the test, including issues that arose, modifications to be made to the approach, suggestions for improvements to CODE, and possible directions for future research.

Chapter 7 TESTING THE APPROACH: LASERS

7.0 INTRODUCTION

Once the definition-oriented approach to concept analysis was developed, it needed to be tested in a practical situation. I chose to create a TKB on the subject field of **lasers** for a number of reasons:

(1) I had been exposed to a number of **laser** concepts when working on optical storage technology subfields and felt that this offered an interesting opportunity to explore a related subject field. Although **lasers** are somewhat different from the concepts investigated for the COGNITERM Project, they are not as different as, for example, concepts from the subject fields of law or economics.

(2) Some preliminary reading showed me that **lasers** were classified in a generic-specific hierarchy, making them suitable for intensional definitions.

(3) Subject-field experts were available at both the University of Ottawa and my current workplace, Bell-Northern Research Ltd. (BNR). Knowing how crucial the role of the expert is to the approach (as well as to traditional terminology work), this factor was one of the most important to the selection of the subject field. While many hours of reading helped me understand most of what is described below, the help of several experts would provide me with missing details. These experts were: Dr. Emory Fortin of the Physics Department of the University of Ottawa, who was kind enough introduce me to the subject field;[75] Dr. Doug Beckett, a laser specialist with the Advanced Optical Devices group of BNR, who was especially helpful in verifying information and helping with decisions on to include in the ITDs; and Dr. Claude Rolland, also a laser specialist with the of BNR Advanced Optical Devices group, who verified the final, natural language ITDs.

It should be remembered that the research framework is knowledge-based, which means that

[75] Dr. Fortin went away for an extended period of time during the test, and was thus consulted only for the first part.

the ITD is not the only conceptual information presented to the user. If there are technical terms in the ITDs (e.g for characteristic names or the name of the generic concept), these terms are explained elsewhere in the TKB via characteristic descriptions or the characteristics of the generic concept. The ITD is a *pointer* to other components of a conceptually rich, well-structured TKB that includes concept relations and attributes, and their individual facets. Thus, while the ITD is an important aid to the user's understanding of the subject field, it is also part of a much greater whole.

7.0.1 The Focus of the Test

Because of the scope of the subject field, I have focused on 16 concepts, most of which are **solid-state lasers**. Records (including ITDs) for these concepts are provided in Appendix B. While most of the information is in English, French equivalents are included where possible.

I looked at several levels of the **laser** hierarchy during this test in order to see whether the difference in hierarchical levels produced different definition problems. I describe the application of the proposed approach for only three levels, namely:

Level 1: **laser**. In order to understand the subject field, I had to understand the topmost concept. Although **laser** is compared to its co-concepts during this test, they are understood to be excluded from the concept system being explored.

Level 2: **solid-state laser, liquid laser**, and **gas laser** (specific concepts of **laser**).

Level 6: **Nd:YAG laser, Nd:YLF laser, Nd:Cr-GSGG laser**, and **Nd:glass laser** (specific concepts of **neodymium laser**).

7.0.2 What is a Laser?

A brief introduction to **lasers** is necessary before the test can be examined. Although **laser** is commonly known to be an acronym for "light amplification by stimulated emission of radiation"

[LAUR86:5],[76] a **laser** (Figure 42) is in fact:

A device that produces a monochromatic, coherent beam of electromagnetic radiation.[77]

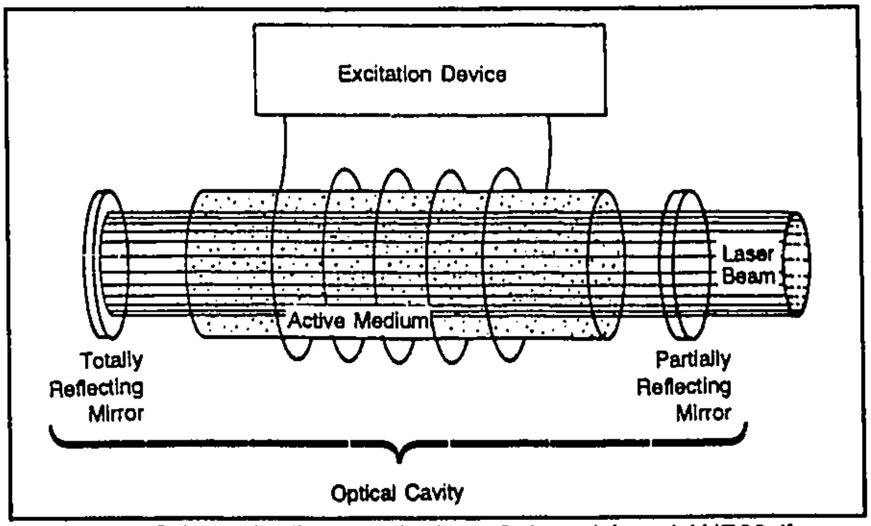


Figure 42. Schematic diagram of a laser [adapted from LAUR86:4].

How this ITD was constructed is explained in the next section. First, however, is an explanation of how a **laser** works. As shown in Figure 42, a **laser** has three main components:

(a) the ***excitation device***, acting as the source of energy;

(b) the ***active medium***, the substance that creates and amplifies the laser light;

(c) the ***optical cavity***, the space containing the active medium, bounded at either end by a mirror.

The excitation device provides energy in the form of photons. Photons can be thought of as packets of waves, in this case, waves of energy. When the active medium absorbs these incoming photons, its atoms become excited and produce *stimulated emission*. Stimulated emission results from a special process by which a photon emitted from an atom comes into contact with another similar

[76] In order to maintain consistency, the documentation codes found in the laser TKB are used here. These reference codes are found in the Bibliographic Reference section of this thesis before their respective references.

[77] This definition resulted from the proposed approach.

atom in an excited state, and then causes (stimulates) another photon *identical to itself* to be emitted. The two mirrors within the **laser** reflect the emitted photons back into the active medium so that they can cause more stimulated emissions. Some photons are allowed to escape through the partially reflecting mirror, however. It is these escaped photons which form the laser beam.

The first **laser**, demonstrated in 1960 by Dr. Theodore H. Maiman, was a **solid-state ruby laser** [LAUR86:17] with an active medium of synthetic ruby crystal. Within three years, **gas lasers** and **liquid lasers** had been developed. These three main types effectively provided the groundwork for the development of all other **lasers**. The subject field of **lasers** continues to evolve today, with scores of different types of **laser** commercially available and still more new ones under development in research labs. As PA91:603 notes about the subfield being explored here:

> The physics and engineering of solid-state lasers is both a mature field and an area burgeoning with new activity. While there are many concepts and laser designs that have been established, each year continues to bring remarkable discoveries that open new avenues of research.

The concepts that were analyzed for the test all belong to one of the more "mature" areas of the subfield: they are well documented and their attributes and relations are firmly established.

7.1 FOLLOWING THE PROPOSED APPROACH

The description of the test follows the same format as the description of the approach itself (cf. chapter 6, Table 9). It is divided into the same five parts, with the individual steps described within each part. Because of the iterative[78] nature of this work, there are times when the same step is repeated for different sets of concepts. This is reflected in the parenthetical information in the section titles.

---

[78] *Iterative* here means "repetitive within a confined number of steps." Although the approach is composed of five main parts, it is likely that the user will have to go back and repeat a given part for different sets of concepts and then continue through with the approach.

7.1.1 Part I: Carrying Out Preparatory Work (All Concepts)

The introductory corpus consisted of two fairly general books on **lasers** (TF87 and LAUR86) which offered a good introduction to the subject field and the theory behind it. The initial challenge was to understand exactly what made the **laser** different from its co-concepts. This required a great deal of reading and note-taking, and resulted in my recording the following information.

According to the documentation, a **laser** is considered a kind of (a) **device** and (b) **source of electromagnetic radiation** (**SER**[79]). The first genus, **device**, seems to arise from an ontological framework, as was discussed in chapter 4 (section 4.2.1). In that framework, a concept system can be divided according to general categories such as **equipment**, **function**, and **process**. The second genus, **SER**, arises from a subject-field framework. Taken together, these genuses imply that a **laser** is a kind of **manufactured SER** (as opposed to a **natural SER**). From the documentation, I created a graph to show the two categories of **SER** and their specific concepts (Figure 43).

From this perspective, then, some[80] of the co-concepts of **laser** are **maser** (a predecessor of the **laser** that produces microwaves), **x-ray machine**, **electric lightbulb**, **fluorescent tube**, and **candle**. All of these devices emit one or more kinds of **electromagnetic radiation** (**ER**). In fact, it is nature of the **ER** emitted by the **laser** that distinguishes it from other **manufactured SERs**. To understand the **laser** and its co-concepts, then, it is essential to understand **electromagnetic radiation**.

Electromagnetic radiation travels in waves; thus, the various kinds of **ER** are distinguished from each other by their different wavelengths. As Figure 44 shows, a wavelength is the distance between two wave peaks (or two wave troughs). Figure 45 shows the different types of **ER**: **radio waves**, **microwaves**, **infra-red**, **visible light** (further divided into **red**, **orange**, **yellow**, **green**, **blue**, **indigo**, and **violet light**), **ultra-violet**, **x-rays**, and **gamma rays**. This range of radiation is known as

[79] This acronym was created for simplicity. The acronym ER is often used for electromagnetic radiation.

[80] For reasons of time, I chose to focus on only a few devices that produce ER.

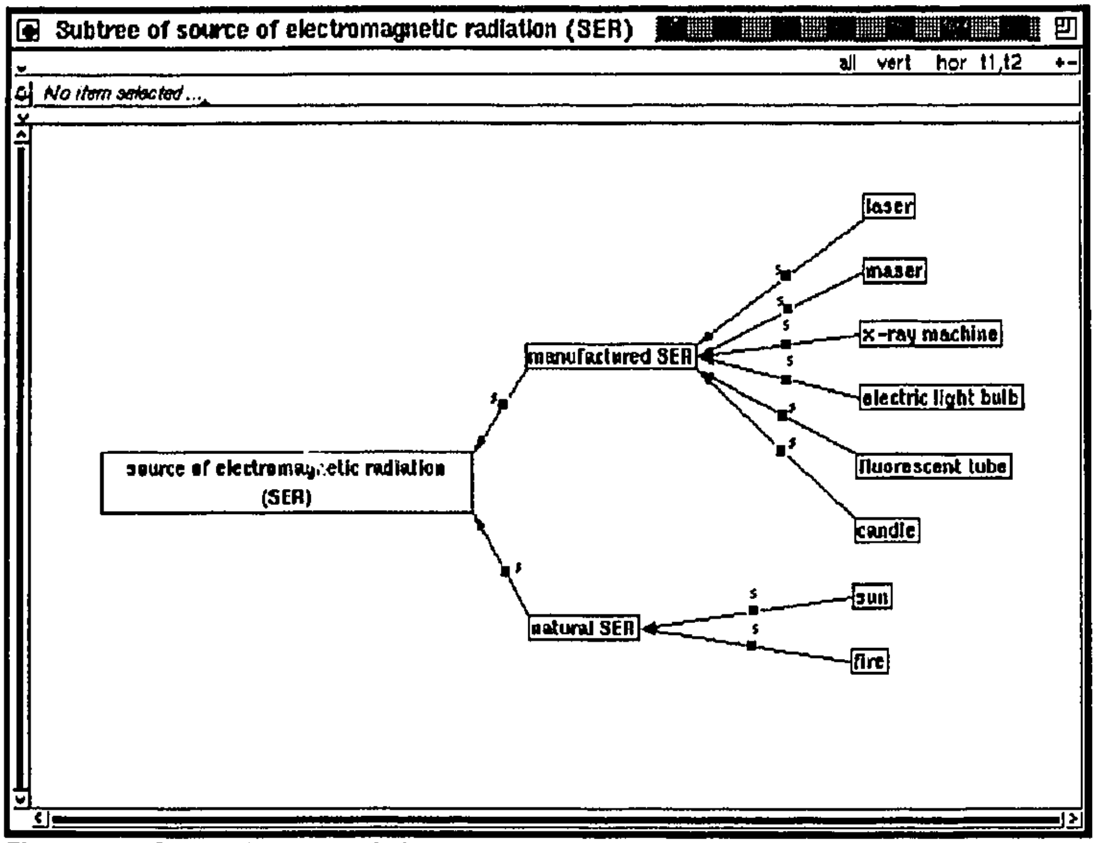


Figure 43. Graph of source of electromagnetic radiation (SER).

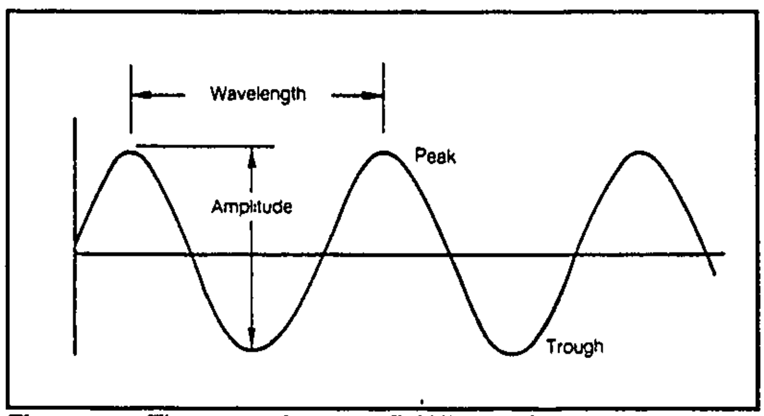


Figure 44. The parts of a wave [LAUR86:24].

the electromagnetic spectrum. The wavelengths of ER vary from less than 0.001 nanometre for gamma rays to hundreds of metres for radio waves. The unit of measurement used when discussing lasers is the typically the nanometre (nm), with lasers producing output from the microwave region to the x-ray region.

Electromagnetic radiation is also often described as a stream of particles, called photons, where each photon can be thought of as a small packet of waves [TF87:5].

Every SER, including lasers, produces electromagnetic radiation. However, according to the

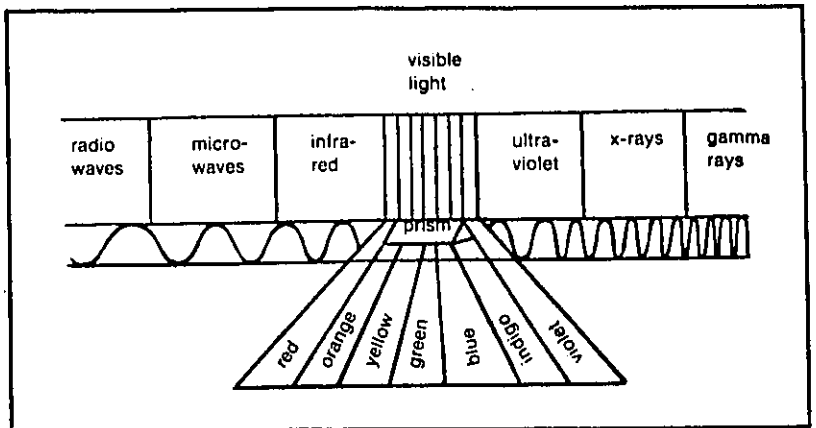


Figure 45. Different kinds of **electromagnetic radiation** (ER) from longest to shortest [ASIMOV90:16].

literature, the output of a **laser** differs from the output of other SERs by four main characteristics: ***chromaticity***, ***coherence***, ***divergence***, and ***power***.

Chromaticity

Chromaticity describes whether the **ER** emitted from the device is composed of one or more wavelengths. A laser beam is *monochromatic*, which means that all the waves in a laser beam are exactly the same wavelength (and thus the same "colour"). There are **lasers** which produce red light, for example, while others produce green light. All other **SERs** produce output which is *polychromatic*. For example, the light given off by an **electric lightbulb** contains all the wavelengths of visible light (appearing to the human eye as white light).

Coherence

Coherence is the relative organization of the ER waves in the output beam; as seen in Figure 46, the waves can be either ***incoherent*** (out of step) or ***coherent*** (in step). It helps to think of waves as marching soldiers: incoherent soldiers all march out of step and move ineffectively as a whole, while coherent ones march in step and thus move more effectively. Prior to the invention of the **laser**, all sources of **ER** were incoherent, which meant that the waves were never in step.

Divergence

Divergence describes how the beam of **ER** spreads out as it travels. Because the **ER** produced by **lasers** is monochromatic and coherent, the beam tends to remain very narrow as it travels over long distances.

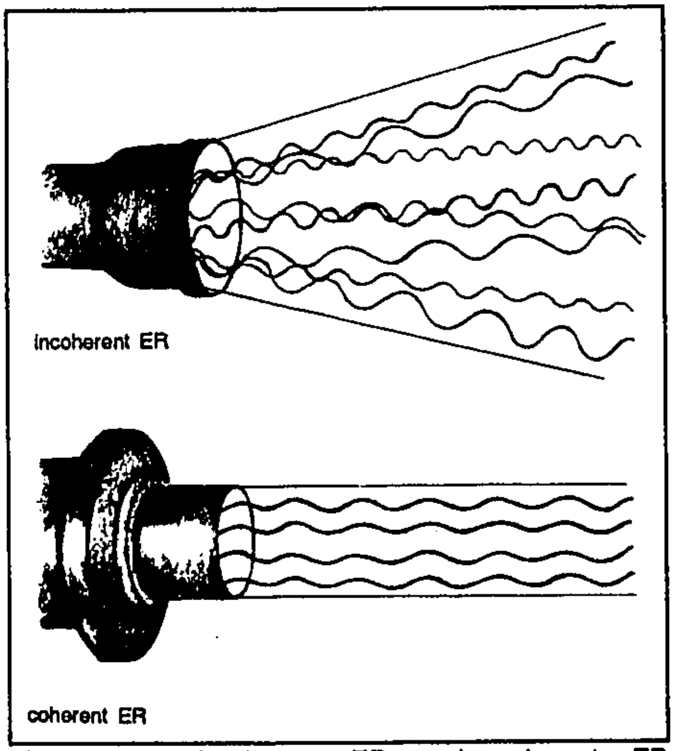


Figure 46. Incoherent ER and coherent ER [ASIMOV90:28].

Power

All of the above three characteristics can cause the **laser's** output to be very powerful: because the output of a **laser** is *monochromatic* (all the waves have the same wavelength) and *coherent* (all the waves are in step), and thus of *low divergence* (spreading out very little over long distances), it can produce high concentrations of energy on very small areas. No other **manufactured SER** is capable of this, as LAUR86:8 notes:

> Until the invention of the laser, there was no way known to science to generate light with these characteristics. This is why the invention of the laser was such an important advance for science and mankind.

By this time, I had read enough preliminary material to understand what a **laser** was and what differentiated it from its co-concepts. I felt confident enough to sketch out a general concept system for the subject field of **lasers** (using information from TF87 and LAUR86) and enter it into CODE (Figure 47).

Being aware of clues in the documentation helped me with this process. For example, TF87 devoted a chapter each to **solid-state lasers**, **gas lasers**, and **dye lasers** (the only type of **liquid laser**

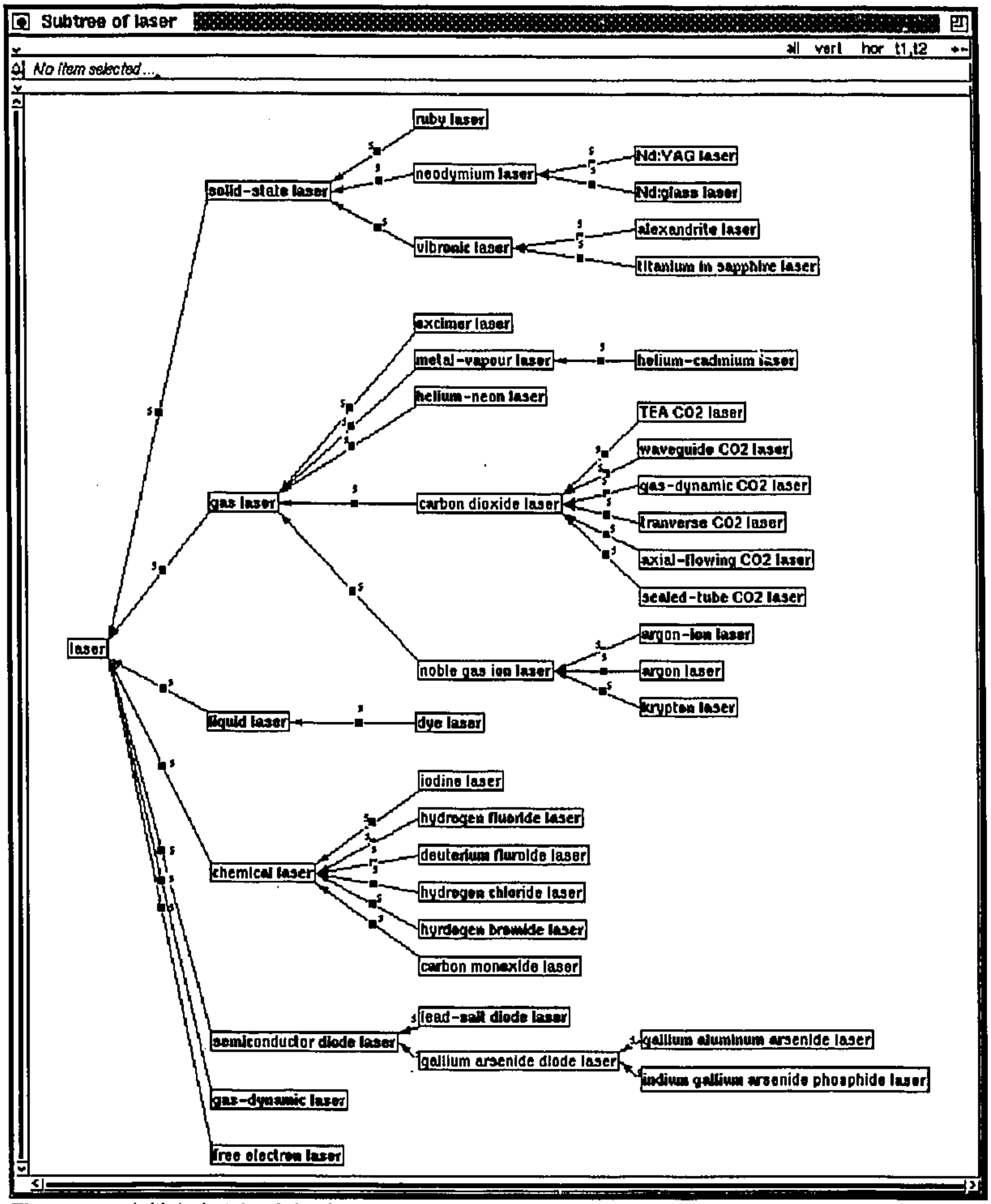


Figure 47. Initial sketch of the **laser** concept system.

discussed in the literature), plus a fourth to "other lasers." Naturally, the three main kinds of **laser** appeared on my graph, forming some of the concept "clusters" seen in Figure 47. Since the documentation implied that the "other lasers" were as important as the three main kinds, they too were given "cluster" status.

I was also keeping track of potential characteristics, recording those which appeared to motivate the subclassification of the subject field. During my preliminary reading, I found that **lasers** were frequently discussed in terms of what the authors explicitly termed their "properties" and "characteristics." For example, LAUR86:6 included a section titled "Special properties of laser light." While at first this made the task of recording characteristics seem almost too simple, it soon became clear that there was an overwhelming number of characteristics, giving me an overwhelming urge to record them *all*. I later discovered that this enthusiasm for recording characteristics could work against me.

Because of all the work done to understand the **laser** and its co-concepts, I had already created a draft template for **laser** based on the gathered information. It made more sense for me to record the information as I did my preliminary research than to wait for the expert to review the **SER** graph. Thus, by the time I had my first interview with the first expert, Dr. Fortin, I had created a sketch of a small, representative concept system of different **SERs** (Figure 43), along with a larger graph of the laser concept system (Figure 47). I also showed him a draft template for **laser**

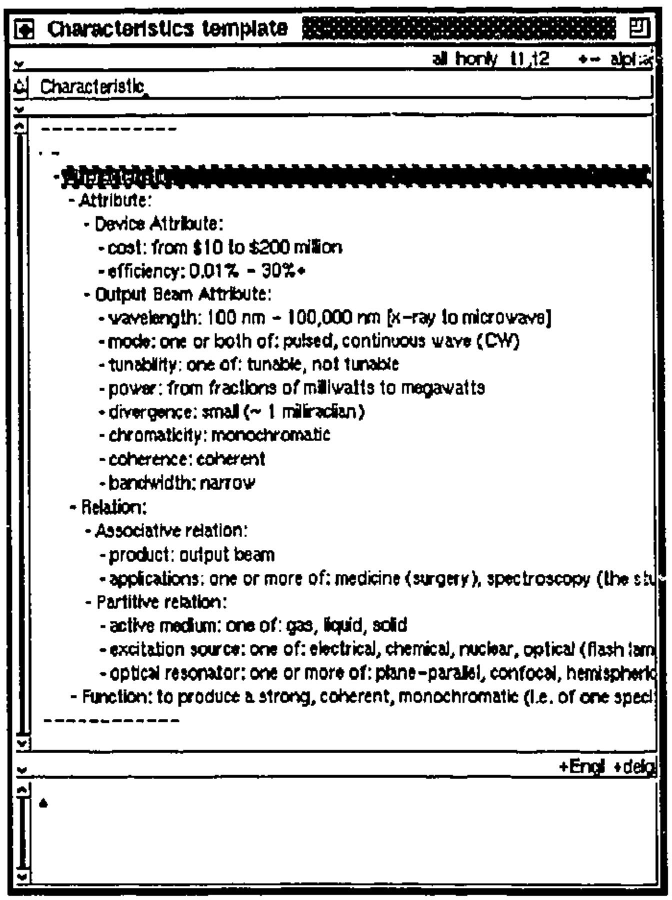


Figure 48. Characteristics template for **laser**.

(Figure 48) and its co-concepts,[81] as well as a list of characteristics gathered for the three main kinds of **laser** (**solid-state laser**, **gas laser**, and **liquid laser**). I presented these to him in paper form, as he was unable to view my work online at the time.[82]

The graphs, as noted in previous chapters, proved helpful in showing Dr. Fortin the work done to that point. The CCMs helped me present the information I had gathered about kinds of **manufactured SER** and kinds of **laser**. It also helped us pinpoint those characteristics which were most important to understanding the various levels of concepts. My notes ensured that I was able to converse with the expert in an intelligent manner, and, I believe, helped me convince him that I was taking the matter seriously and did not intend for him to do all the work. (I had had some initial difficulty convincing Dr. Fortin to act as my expert, perhaps because of my non-scientific background, and I believe I won him over at this point.)

The main points that resulted from this initial interview were the following:

— the principal characteristics that differentiate the **laser** from other **manufactured SERs** are *coherence*, *chromaticity*, *divergence*, and *power*

— Dr. Fortin suggested I add the concept **unmodulated radio emitter** to the graph (to represent a device that emits radio waves)

— the main components of a **laser** are the *active medium*, the *excitation device*, and the *optical resonator*; these could also be considered the most important characteristics for kinds of **laser**

— there are many ways of classifying the **laser** subject field; a true graphic representation would result in "a lot of criss-crossing" [FORTIN93] between the concepts (i.e. it would be multidimensional)

---

[81] Although only the template for **laser** is shown, the initial template for the co-concepts of **laser** will look the same.

[82] It should be noted at this point that the ideal situation, with the expert by my side at the computer, never materialized during this test: I was forced by circumstances to always present my information in paper form to the experts. On the one hand, I am not sure that the process would have gone much faster if I had done it "live." On the other hand, it would have been more interesting for the expert to have been able to browse through the TKB rather than look at the information on paper.

— if we had to restrict ourselves to one dimension,[83] it should be based on the ***active medium***, as this is the most common way of classifying the subject field; thus the main divisions were restricted to **solid-state laser**, **liquid laser**, and **gas laser**, with the other concepts placed under those three

— I would concentrate on **solid-state lasers**, as this was Dr. Fortin's area of expertise; he was also able to recommend another expert who specialized in this subfield

After annotating the various characteristics and concepts according to the expert's comments, I selected the principal documentation, including a range of books and journals on **lasers** (some specifically dealing with **solid-state lasers**, but others more general, for the co-concepts of **solid-state laser**) and an encyclopedia of lasers. The bibliographic information was entered in CODE and the reference codes selected (these are found in the Bibliographic References of this thesis).

The final step of this first part was to read the subfield documentation and establish a preliminary concept system for the subfield. The results of this work can be seen in Figure 49. The first meeting with Dr. Fortin helped clarify some of the generic-specific relations (e.g. the three main kinds of **laser**, and several main kinds of **solid-state laser**). This was annotated in the background information of the appropriate concepts.

At this stage, I realized that, in order to fully understand the research I was going to do on **solid-state lasers**, it would be better to define the concept of **laser** as soon as possible. Thus, I set

[83] This strategy of adhering to one dimension brings up several issues: (1) The terminologist must strive to be faithful to her selected dimension. Although she may be taking a unidimensional approach (dividing each subhierarchy according to only one characteristic), she must keep in mind that different authors may work within different dimensions, or that they may take a multidimensional approach. (2) Restricting the subject field to one dimension may mean that certain concepts are not represented. For example, some **lasers** are differentiated from others not by the ***active medium*** but by the ***excitation source***. **Chemical lasers** are defined thus: "The term chemical laser refers not to the state of the [active] medium, but to the method of creating a population inversion [exciting the atoms]. Although the chemical can be in the solid, liquid or gaseous state, most chemical lasers use gases as the [active] medium..." [OSHEA77:142]. It is possible that, within the dimension of ***active medium***, some **chemical lasers** may not be represented.

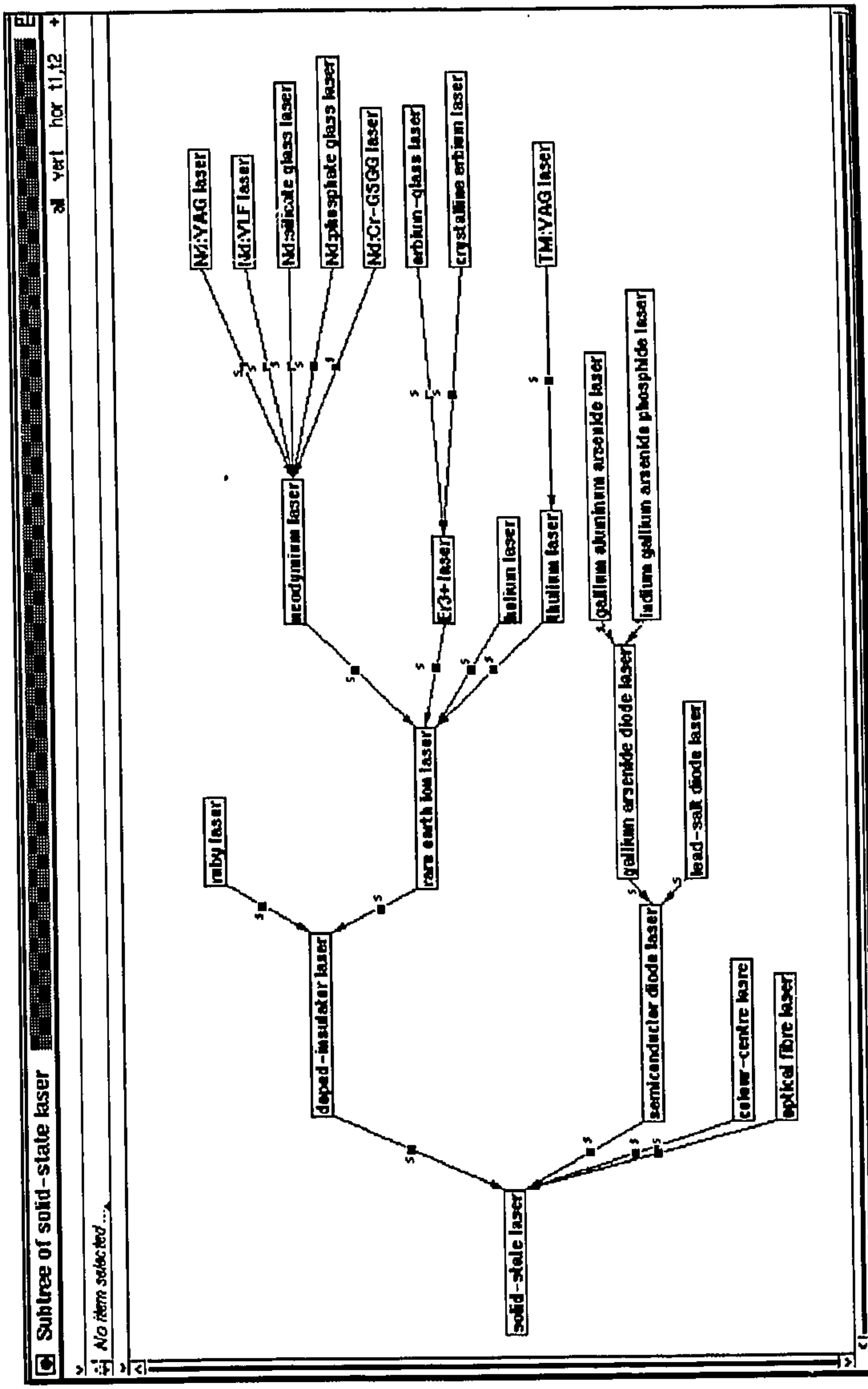


Figure 49. Subfield graph for solid-state lasers after first interview with expert.

aside the **solid-state laser** research and moved ahead in the work on **laser**.

7.1.2 Part II: Establishing the Characteristics Template (Kinds of Manufactured ER)

Since the template for **laser** had already been created and verified with Dr. Fortin, I could begin scanning.

7.1.3 Part III: Carrying Out Primary Knowledge Acquisition and Revision (Kinds of Manufactured ER)

Since so much work had gone into the preliminary research, there was little scanning to do for **laser** and its co-concepts.

7.1.4 Part IV: Constructing the Draft ITD (Laser)

I had created the template by exploring only the characteristics that appear to distinguish a **laser** from its co-concepts. Upon closer inspection of the literature, however, two of the differentiating characteristics (shown in the CCM in Figure 50) appeared typical:

Matrix of laser and coordinates

	manufactured SER: (s)						
	laser	candle	electric light bulb	x-ray machine	unmodulated radi	fluorescent tube	maser
coherence	coherent	incoherent	incoherent	incoherent	incoherent	incoherent	incoherent
chromaticity	monochromatic	polychromatic	polychromatic	polychromatic	polychromatic	polychromatic	polychromatic
divergence	low	high	high	high	high	high	high
power density	high	low	low	low	low	low	low

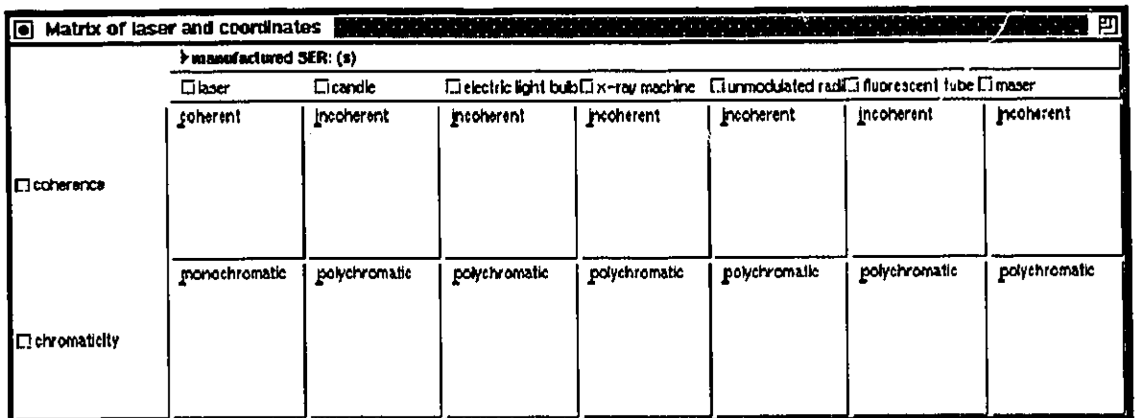
Figure 50. CCM of kinds of **manufactured SER** with differentiating characteristics showing.

(1) *power:high* is typical because not all lasers are in fact powerful. Although there are **lasers** which can produce very powerful beams (more than 2 million watts), there are also **lasers** which have

quite feeble beams, such as those used to read supermarket bar codes. "Their power is less than a thousandth of a watt, thousands of times weaker than the smallest nightlight" [ASIMOV90:47].

(2) ***divergence:low*** is also typical, since there are lasers, such as the **semiconductor lasers** found in CD players and CD-ROM drives, whose beams actually do spread out quite a bit over distance. (This is not a problem in CD players and CD-ROM drives because their laser beams cover very small distances and thus do not get a chance to spread out much.) While **lasers** can be concentrated and made to come to a very fine point (because they are coherent), they do not all naturally have the characteristic ***divergence:low***.

Thus, ***power:high*** and ***divergence:low*** were masked out, with the resulting CCM shown in Figure 51.



Matrix of laser and coordinates

	manufactured SER: (s)						
	laser	candle	electric light bulb	x-ray machine	unmodulated radi	fluorescent tube	maser
coherence	coherent	incoherent	incoherent	incoherent	incoherent	incoherent	incoherent
chromaticity	monochromatic	polychromatic	polychromatic	polychromatic	polychromatic	polychromatic	polychromatic

Figure 51. CCM for kinds of **manufactured SER** with n+s characteristics showing.

Feeling fairly confident of my knowledge at this point, I decided to select the n+s characteristics and then have the expert review the draft ITD (using the "tout le défini" questions as guides). From what I had learned, I felt that ***coherence:coherent*** and ***chromaticity:monochromatic*** were both n+s characteristics. As ASIMOV90:38 noted when comparing lasers with masers, "A maser[84] produces a coherent monochromatic beam of microwaves; a laser produces a coherent monochromatic beam of light.[85]"

Thus, I constructed the draft ITD for **laser** from the CCM in Figure 51, which was:

> *A manufactured source of electromagnetic radiation which produces a monochromatic, coherent beam.*

7.1.5 Part V: Testing the Draft ITD (Laser)

Since I wanted to verify that the characteristics in my ITD were both necessary and sufficient, I tested it out with the second expert (Dr. Beckett) by asking the first "tout le défini" question:

> Are all concepts which have the generic concept (B) and the set of characteristics (C) denoted by the term (A)?

or, in other words:

> Are all **manufactured SERs** that produce a *monochromatic, coherent beam* called **laser**?

The answer was *yes*. In other words, the expert could not imagine any **manufactured SER** that had all the characteristics in the ITD that would not be called a **laser**.

For the sake of argument, some points that could have been considered when combining the n+s characteristics are presented below:

— If I had included *power:high* with the other two n+s characteristics, the ITD would have excluded some kinds of **semiconductor laser** which have weak beams. Thus, the ITD would have been too narrow.

---

[84] This question of **masers** was troubling at first, since **masers** and **lasers** are often represented in the literature as co-concepts. Dr. Beckett explained that a **maser** was in fact a kind of **laser**. The **maser** (acronym for *Microwave Amplification by Stimulated Emission of Radiation*) was invented in 1953 by Dr. Charles H. Townes—before the **laser**. The **optical maser** followed in 1960 and quickly became known as the **laser**. Once this became clear, I modified my **SER** graph so that **maser** became a specific concept of **laser**.

[85] Although the output of a laser is sometimes referred to as "light" in the more general literature, the correct term is "electromagnetic radiation."

— If I had included the characteristic *divergence:low*, the ITD would have excluded **semiconductor lasers**, which have fairly a high divergence, according to Dr. Beckett. Thus, the ITD would have been too narrow.

While writing the final, natural language ITD, I decided that the genus **manufactured SER** seemed rather artificial. I chose to use **device** as the genus and imply that a **laser** was a "source of ER" by specifying the type of beam it produced (many defining contexts in the literature had done this; cf. Appendix B, **laser**, Linguistic Information). With these modifications, the final natural language ITD for **laser** was

A device that produces a monochromatic, coherent beam of electromagnetic radiation.

7.1.6 Part II: Establishing the Characteristics Template (Kinds of Laser)

It was now time to create a template for the second level of **laser** concepts (**gas laser, liquid laser**, and **solid-state laser**). Unfortunately, it was at this point that my enthusiasm for collecting characteristics got the better of me. The draft template included the three main characteristics noted by Dr. Fortin and me, along with others that had often appeared in the documentation:

— ***active medium*** (with its components, ***host material*** and ***active species***)

— ***excitation source***

— ***optical resonator***

— ***commercial value***

— ***wavelength***

— ***mode***

— ***applications***

The set of characteristics which differentiated the output of **lasers** from that of other **manufactured SERs** (***coherence***, ***chromaticity***, ***divergence***, and ***power***) was also included, since they had been inherited from **laser**.

In retrospect, the initial three characteristics that Dr. Fortin and I had agreed upon during our first meeting would have been sufficient for a "kinds of laser" template. While many of the characteristics above were often mentioned in the documentation, in retrospect, I should have been more selective in the characteristics included in the template. I should have also masked out the four laser output characteristics at this point. They had fulfilled their purpose in distinguishing **lasers** from other kinds of **manufactured SER** and would not be needed for future knowledge acquisition.

During the second meeting with Dr. Fortin, he verified the state of the **solid-state laser** graph, at which point (on reflection) we got a bit off track. Forgetting our focus on ***active medium*** as the main "classification" characteristic, we started discussing **tunable lasers.**[86] Dr. Fortin suggested that I include **tunable solid-state lasers** along with **doped-insulator lasers** and **semiconductor diode lasers.** Although I would still concentrate on **neodymium lasers**, this new development led to the characteristic ***tunability*** being added to the already overburdened characteristics template. Dr. Fortin also suggested ***commercial value*** as an interesting characteristic.

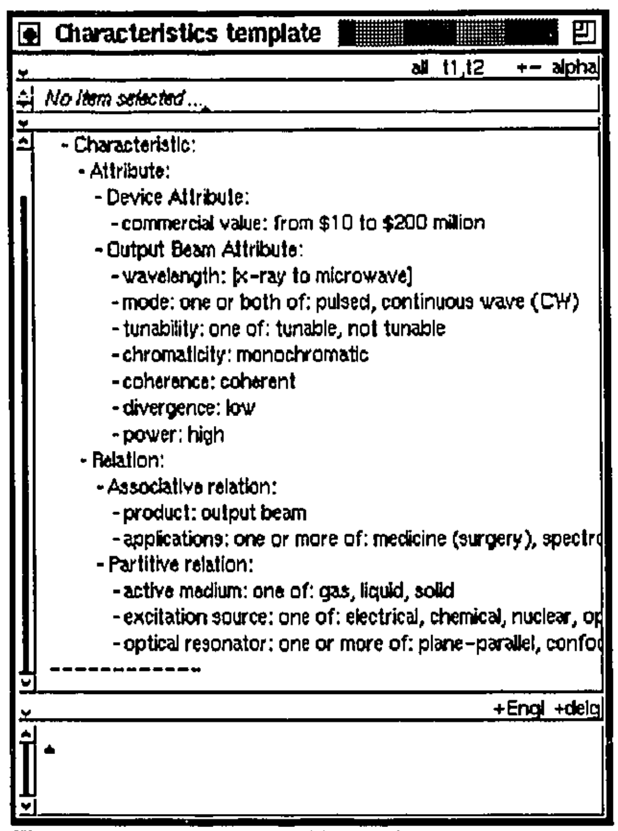


Figure 52. Template for kinds of **laser.**

Since Dr. Fortin agreed to the overall content of the template and the general outlay of the graph, I recorded his code (*FORTIN93*) in the appropriate ***EXPERT*** facets. The resulting template and graph are shown in Figure 52 and Figure 53.

[86] A **tunable laser** can produce beams over a number of wavelengths. Except for **gas lasers** and some **solid-state lasers**, most **lasers** can produce a beam of only one wavelength.

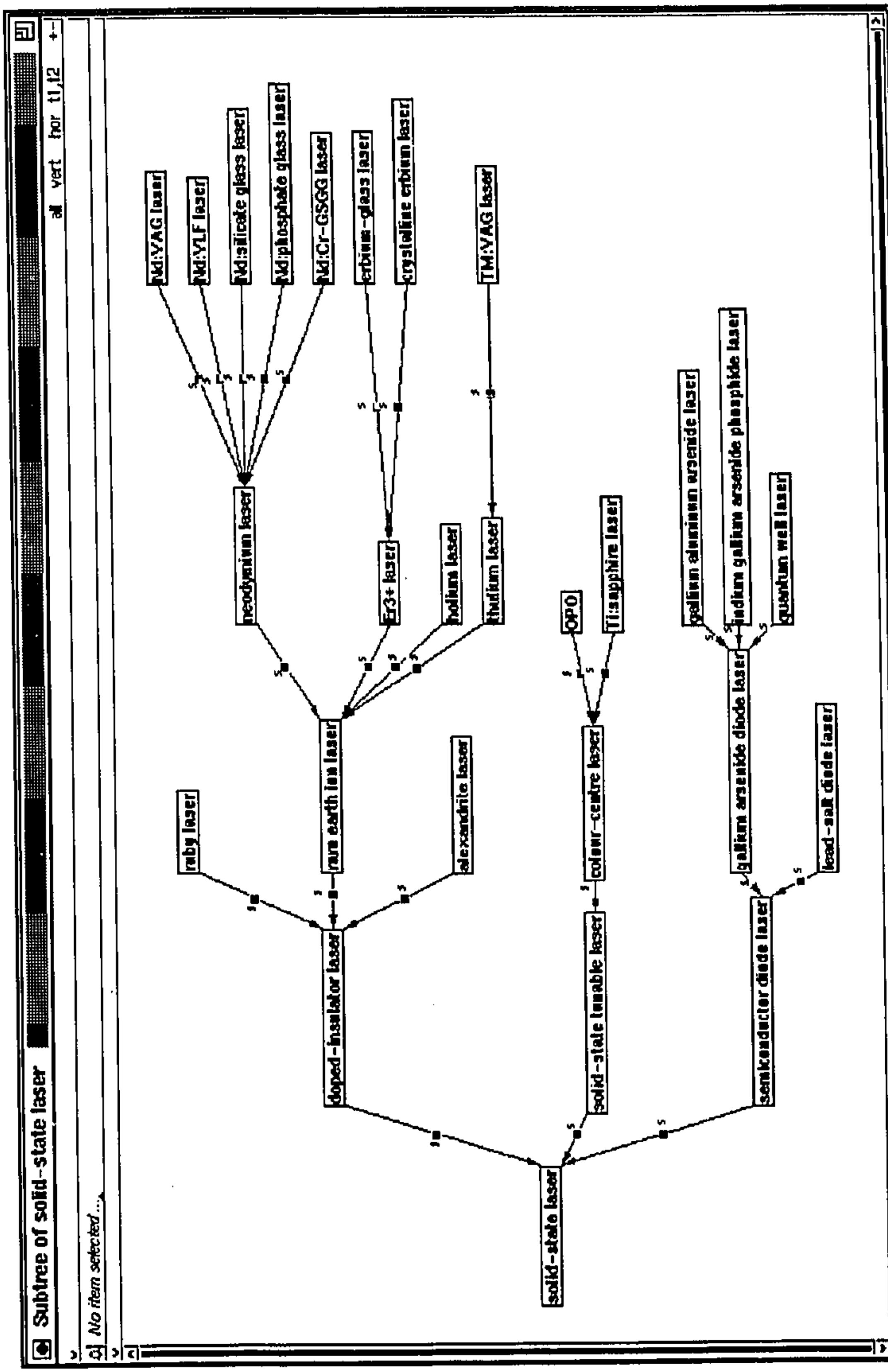


Figure 53. Subfield graph for solid-state lasers after second interview with expert.

In retrospect, I did not fulfil my job as a "quality (or quantity) control" manager. I should have kept in mind that we had agreed on a particular dimension (which would have stopped ***tunability*** from becoming an issue), and I should have been more critical of the rather large number of characteristics in my template.

7.1.7 Part III: Carrying out Primary Knowledge Acquisition and Revision (Kinds of Laser and Kinds of Neodymium Laser)

With my template and preliminary concept system in place, I was ready to begin scanning. Although I generally scanned using a top-down approach, I often found it more practical to record information whenever it was found in the documentation, regardless of the level I was working on, in order not to "lose" information.

Modifying the template

Since I was working at a variety of hierarchical levels, I knew it was possible that the initial template would have to be modified as I moved through the hierarchy. However, rather than the template changing, it was the focus of the template that changed. Initially it was ***active medium*** which appeared to be the most important characteristic (at the second and third levels). At the fourth and fifth levels (kinds of **doped-insulator laser** and kinds of **rare earth ion laser**), it was ***active species***, a component of ***active medium***, which became the most important characteristic. At the sixth level (kinds of **neodymium laser**), it was ***host material***, also a component of ***active medium***, which appeared to distinguish the different kinds of **neodymium laser**. Thus, although the template did not have to be changed during scanning, my perception of which characteristics were important did.

Looking for clues

Knowing that I should look for "clues" in the documentation and annotating characteristics helped me with my scanning: key sentences, chapter and section titles, and tables of properties (which were quite common in the laser literature) would jump out from the page. Some examples:

1) HECHT92:91 efficiently provides what he feels to be the two most important characteristics: "There are many potential criteria for classifying lasers, but the two most useful ones are the *type of active medium* and the *way in which it is excited* [my emphasis]."

2) TF87:30 includes a rather extensive table called "Laser properties and applications," which divides 15 different kinds of **laser** into the three main kinds (solid, liquid, and gas), and provides five different properties for each (wavelength, excitation source, pulsed or cw [mode], power, and application).

3) LUPA92:140 provides an overview of the entire concept system, noting:

> Laser types can be categorized in a number of ways. In general, there are gas, liquid, and solid lasers. Subclassification of neutral gas, ion, metal vapor and molecular lasers exist in the category of gas lasers. The only liquid laser discussed here is the dye laser. Solid lasers may be crystalline, amorphous (at least in the case of the Nd–glass laser), or even a semiconductor.

Sometimes, however, the clues were less easy to spot, because the authors did not always spell out the important details. For example, there were some nagging blanks in the TKB's generic–specific relations, even towards the end of scanning. This was especially true with the **ruby laser**: everyone referred to it as a kind of **solid–state laser**, but ignored the fact that there may be other concepts between the **ruby laser** and the **solid–state laser**. Since I had ascertained that the two main kinds of **solid–state laser** were **doped–insulator laser** and **semiconductor laser**, I at least knew that the **ruby laser** was a kind of **doped–insulator laser** (since it is not made of semiconductor material). But the two main kinds of **doped–insulator laser** were **rare earth ion laser** and **transition metal laser**. Was **ruby laser** one of these? Or was it off by itself with no co–concepts (something I instinctively found hard to believe)?

Most authors were unhelpful in describing the exact genus of the **ruby laser:**

1) PA91:603: "This first system was a solid-state laser; a ruby crystal served as the active element and it was pumped with a flash lamp."

2) MESS90:81: "The first laser invented – ruby – is also solid state."

3) SILF91:221: places ruby within a section called "Solid-state lasers" with the subsections titled 1. Neodymium Lasers, 2. Ruby Laser, and 3. Color Center Lasers, implying that these three are co-concepts (which I knew from my research to be untrue).

I finally found the answer when I had all but finished scanning:

4) HECHT92:96 implies that the **ruby laser** is a kind of **solid-state laser** in a general section on **solid-state lasers**. Unfortunately, in a chapter devoted to the **ruby laser**, he does not say what kind it is. He does, however, mention that "ruby laser rods (i.e. the active medium) are grown from sapphire (Al_2O_3) doped with about 0.01 to 0.5 percent chromium from a synthetic ruby crystal, colored red or pink with about 10^{19} chromium atoms per cubic centimetre" [HECHT92:419]. He also provides a small graph titled "Energy levels of *Cr^{3+}* [my emphasis] in ruby, showing pump and laser transitions" [HECHT92:420].

It was only from this series of clues, culminating in the graph title mentioning Cr^{3+}, combined with the subfield representations of OSHEA77 (who calls "non-semiconductor solid-state lasers" **doped-insulator lasers**), and PA91 (who divides **solid-state lasers** into **rare earth ion lasers** and **transition metal lasers**), that I finally determined that the **ruby laser** was a kind of **Cr^{3+} laser** (see Figure 54), itself a kind of **transition metal laser**. This in turn was a kind of **doped-insulator laser,** which was a kind of **solid-state laser**.

Using the CCM

The CCM did not prove to be as useful during scanning as I thought it would be, primarily because it cannot have a Conceptual Facets browser attached to it. Without attached facets, it was impractical to input information into the CCM, considering the emphasis being put on annotating the information.

It was towards the end of scanning, during the information revision, that the CCM proved valuable for carrying out quality control. It allowed me to quickly compare and standardize the form and content of the characteristics of co-concepts and to see where more research was needed. It also allowed me to present the information in a coherent manner to the expert, albeit without attached facets.[87] However, the fact that I had been forced to monitor the information so closely via the facets made me confident of its quality. Had I simply recorded the information without the facets acting as a kind of "watchdog," I may have been less meticulous. Having to justify my decisions made me more careful of how I arrived at those decisions.

When I was almost done scanning, I consulted Dr. Beckett about the state of my information. Using printouts of the **solid-state laser** subgraph and the CCMs, I was able to quickly explain the work done to that point. Dr. Beckett suggested a few changes to the subgraph, specifically the removal of the **tunable solid-state laser** subfield, because it did not fit within my dimension of ***active medium***. If I had been dividing my subfield up according to ***tunability***, or if I were including both dimensions, then that particular subfield would have been appropriate. Dr. Beckett also verified the recorded characteristics, making a few suggestions regarding characteristic values (e.g. removing *nuclear* from the value of ***excitation source*** for **laser** and not specifying that the ***active medium*** of the **semiconductor diode laser** is *2 semiconductor crystals*, since it can in fact be one or more crystals). The final graph for the **solid-state laser** subfield is found in Figure 54.

[87] While I could present the facets in a printout of the concept information, this meant switching back and forth between two different documents, which made examining the facets rather difficult.

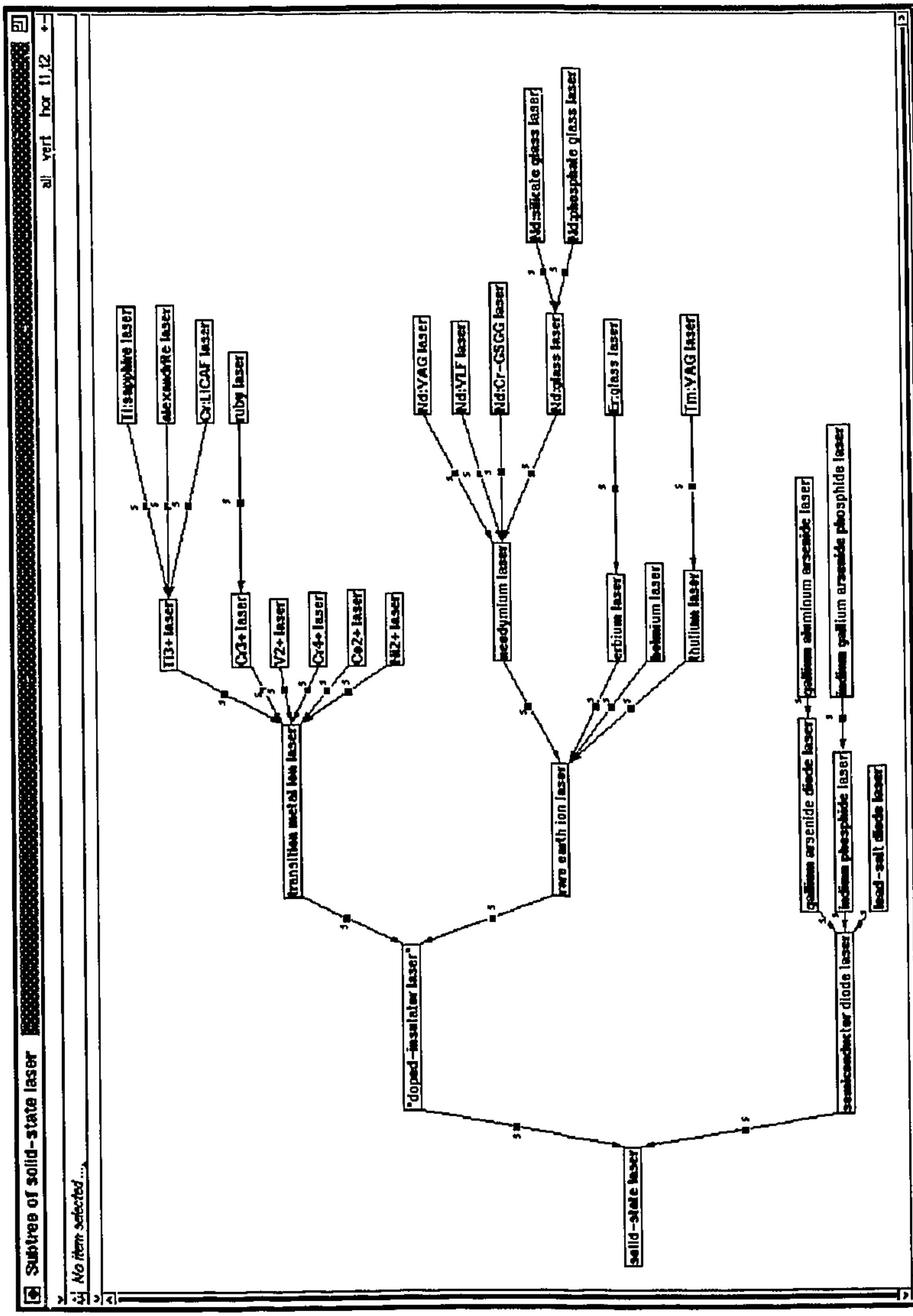


Figure 54. Final graph of solid-state laser subfield.

At this stage, I had to finalize the modality of the various characteristics of the two levels (kinds of **laser** and kinds of **neodymium laser**), which I planned to confirm with Dr. Beckett at our next meeting. I put the CCM in differentiating mode, since, for definition construction, I was only interested in those characteristics which differentiate co-concepts.

Keeping in mind the test for a necessary characteristic ("for a characteristic (C) to be necessary to a concept (X), we must be able to say that 'X must have/be C'"), I found that the easiest way to establish the potential modality of each characteristic was to ask myself the question: "Do ALL instances of this concept have this characteristic?" Most of the time I could confidently answer the question and enter a value of *N(ECESSARY)* for "yes" or *T(YPICAL)* for "no" in the characteristic's *MODALITY* facet. When unsure of the answer, I recorded *?(UNDECIDED)* and made a note to ask the expert his opinion.

The answers to the modality questions for the second-level **laser** concepts (**solid-state laser**, **liquid laser**, and **gas laser**), asked while consulting Figure 55, are found below.

A. Solid-state lasers

Necessary

1. ***active medium:*** all solid-state lasers have a solid active medium

Typical

1. ***excitation source:*** all solid-state lasers have an electrical or optical excitation source, but only one at a time
2. ***tunability:*** all solid-state lasers are either tunable or not
3. ***power:*** all solid-state lasers have a power *within* that range
4. ***wavelength:*** all solid-state lasers produce a wavelength *within* that range, but not with that range

? (Undecided = Questions for expert)[88]

1. ***optical resonator:***

Q: one or two documents say that some **lasers** don't have to have an optical resonator: which **lasers** (of gas, liquid, solid) don't?

A: Gas. However, since a given **solid-state laser** has only one of these values at a time, it is considered typical.

[88] The answers to these questions have been included at this point to avoid confusion.

Matrix of statements of solid-state laser and coordinates

	laser: (s)		
	solid-state laser	gas laser	liquid laser
active medium	solid (one of: impurity-doped insulator, semiconductor)	gas	liquid
excitation source	one of: electrical, optical (flash lamp, another laser)	one of: electrical, chemical, optical (flash lamp, another laser)	optical (flash lamp, another laser)
tunability	one of: tunable, not tunable	one of: tunable, not tunable	tunable
optical resonator	one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)	one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable), none	one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
power	from fractions of milliwatts to 1 terawatt	up to 10 megawatts	from fractions of milliwatts to megawatts
wavelength	[x-ray to microwave]	[x-ray to far infra-red]	[gamma ray to infra-red]

Figure 55. CCM for kinds of **laser** with differentiating characteristics showing.

2. ***host material:***

Q: ***Host material*** has been deemed "typical" in some documents (i.e doesn't exist for **semiconductor lasers,** nor for **ion lasers** such as **argon laser** and **krypton laser**). Is this true?

A: Technically yes. This is where we realized that only **doped-insulator lasers** should really have their ***active medium*** divided into two components (i.e. ***host material*** and ***active species***); thus these two characteristics had be removed from other kinds of **laser** and initiated at the level of the **doped-insulator laser**.

3. ***active species:***

Q: Is ***active species*** then also typical? or does it equal the ***active medium*** if there is no ***host***?

A: If there is no ***host***, then ***active species*** and ***active medium*** can be considered the same thing (see 2).

B. Gas laser

Necessary

1. ***active medium:*** all gas lasers have a solid active medium

Typical

1. ***excitation source:*** all gas lasers have an excitation source, but it can be only one of those values at a time
2. ***tunability:*** all gas lasers are either tunable or not, but not both
3. ***optical resonator:*** not all gas lasers have an optical resonator
4. ***power:*** all gas lasers have a power *within* that range of power, but not with that range of power
5. ***wavelength:*** all gas lasers have a wavelength *within* that range, but not with that range

C. Liquid laser

Necessary

1. ***active medium:*** all liquid lasers have a liquid active medium
2. ***excitation source:*** all liquid lasers have an optical excitation source
3. ***tunability:*** all liquid lasers are tunable

Typical

1. ***optical resonator:*** all liquid lasers have an optical resonator, but it can be only one of those values at a time
2. ***power:*** all liquid lasers have a power within that range, but not with that range
3. ***wavelength:*** all liquid lasers have a wavelength within that range, but not with that range

The answers to the modality questions for the sixth-level laser concepts (**Nd:YAG laser, Nd:YLF laser, Nd:glass laser,** and **Nd:Cr-GSGG laser**), using the CCM in Figure 56, are found below.

D. Nd:YAG laser

Necessary

1. ***host material:*** all Nd:YAG lasers have yttrium aluminum garnet (YAG) as their host material
2. ***sensitizer:*** no Nd:YAG lasers have any sensitizer in their active species

Typical

1. ***power:*** all Nd:YAG lasers have a power within that range, but not with that range
2. ***wavelength:*** not all Nd:YAG lasers produce beams with all three of those wavelengths

E. Nd:YLF laser

Necessary

1. ***host material:*** all Nd:YLF lasers have yttrium lithium fluoride (YAG) as their host material
2. ***sensitizer:*** no Nd:YLF lasers have any sensitizer in their active species

Matrix of statements of Nd:YAG laser and coordinates

	neodymium laser: (s)			
	Nd:YAG laser	Nd:Cr-GSGG laser	Nd:YLF laser	Nd:glass laser
host material (part)	crystal: yttrium aluminum garnet (YAG)	crystal: gadolinium scandium gallium garnet	crystal: yttrium lithium fluoride (YLF)	glass (one of:phosphate, silicate, fluorophosphate)
sensitizer (part)	none	Cr3+	none	none
power	from a few milliwatts to several megawatts	from fractions of milliwatts to 1 terawatt	from fractions of milliwatts to 1 terawatt	up to 1 terawatt
wavelength	946 nm, 1064 nm, 1319 nm	1061 nm, 1335 nm	1047 nm, 1053 nm, 1313 nm	1059 nm

Figure 56. CCM for kinds of **neodymium laser** with differentiating characteristics showing.

Typical

1. ***power:*** all Nd:YLF lasers have a power within that range, but not with that range
2. ***wavelength:*** not all Nd:YLF lasers produce beams of both those wavelengths

F. Nd:glass laser

Necessary

1. ***host material:*** all Nd:glass lasers have glass as their host material
2. ***sensitizer:*** no Nd:glass lasers have any sensitizer in their active species
3. ***wavelength:*** all Nd:glass lasers produce beams with that wavelength

Typical

1. ***power:*** all Nd:glass lasers have a power within that range, but not with that range

G. Nd:Cr–GSGG laser

Necessary

1. ***host material:*** all Nd:Cr–GSGG lasers have Cr–GSGG as their host material
2. ***sensitizer:*** all Nd:Cr–GSGG lasers have Cr^{3+} acting as a sensitizer (to the active species, Nd^{3+})

Typical

1. ***power:*** not all Nd:Cr–GSGG lasers have a power within that range

2. ***wavelength:*** not all Nd:Cr–GSGG lasers produce both these wavelengths

I was now ready to finalize the characteristics' modality with the expert; I would select the potential n+s characteristics with his help at the same time. For the most part, the expert agreed with my decisions, and quickly provided the few answers needed for the questions cited above.

7.1.8 Part IV: Constructing the Draft ITDs (Kinds of Laser and Kinds of Neodymium Laser)

With so few characteristics to compare, the process of constructing the draft ITDs took relatively little time.

Kinds of laser

For the **solid–state laser**, **gas laser**, and **liquid laser** co–concepts, the CCM now showed only ***active medium***, ***excitation source***, and ***tunability*** (since the latter two were necessary characteristics for **liquid laser**). This is where art and subject–field expertise came into play. When I asked Dr. Beckett if he felt that ***excitation source*** was essential to understanding any of the three concepts, he said "absolutely not."[89] We also decided that ***tunability*** should be removed because we were not working within that dimension. Thus, those two non–n+s characteristics were removed, leaving only ***active medium***. And there (shown in Figure 57), with the genus at the top of the CCM and the differentia in the value panes, were the building blocks for the draft ITDs of **solid–state laser, gas laser**, and **liquid laser**.

Kinds of neodymium laser

The process of selecting the n+s characteristics for the four kinds of **neodymium laser** was

[89] My own initial instinct had been to say "no" as well; the reason may be that there is still a selection of more than one value, and thus it doesn't seem as though the characteristic is "ready" to be considered n+s.

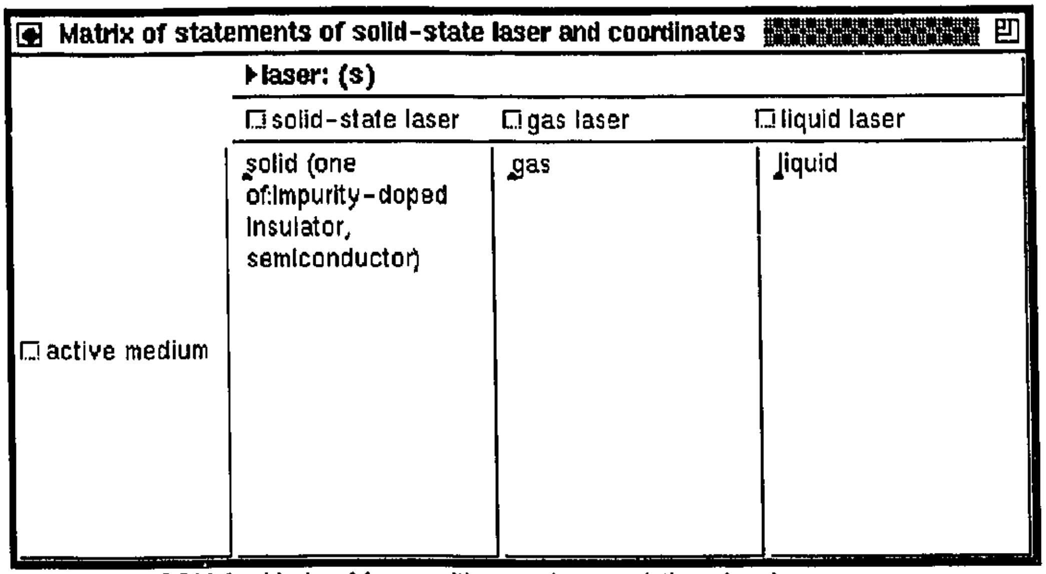


Figure 57. CCM for kinds of **laser** with n+s characteristics showing.

similar to that used for the second level concepts, with one small difficulty. Again, with only differentiating necessary characteristics showing, few characteristics were displayed on the CCM: only ***host material*** and ***sensitizer***, in fact, as shown in Figure 58. ***Host material*** obviously differentiated the four kinds of **neodymium laser**, and Dr. Beckett felt it should be included in the ITD of each concept. It was ***sensitizer*** which initially posed a problem, for while each concept essentially has the same ***active species*** (*Nd^{3+}*), the **Nd:Cr–GSGG laser** also has an element Cr^{3+} added to the Nd^{3+} to improve its ability to produce stimulated emission. Initially I thought of including ***active species*** in each ITD along with ***sensitizer*** (breaking all my carefully made rules); however, this would make the ITDs of the four concepts redundant, since the ITD of their genus, **neodymium laser**, is *A rare earth ion laser whose active species is Nd^{3+}*. To include ***active species*** in an ITD for **Nd:YAG laser**, for example, would result in the following ITD:[90]

> *A [rare earth ion laser whose active species is Nd^{3+}] whose host material is yttrium aluminum garnet (YAG) and whose active species is Nd^{3+} with no sensitizer.*

[90] The genus, **neodymium laser**, is replaced with its ITD in square brackets to better show the redundancy.

Matrix of Nd:YAG laser and coordinates

	neodymium laser: (s)			
	Nd:YAG laser	Nd:Cr-GSGG las	Nd:YLF laser	Nd:glass laser
host material (part)	crystal: yttrium aluminum garnet (YAG)	crystal: garnet	crystal: yttrium lithium fluoride (YLF)	glass (one of:phosphate, silicate, fluorophosphate)
sensitizer (part)	none	Cr3+	none	none

Figure 58. CCM for kinds of **neodymium laser** with n+s characteristics showing.

Not only would the ITD have been redundant, it would have described something that the concept did *not* have (a *sensitizer*), which did not seem justified. Thus, I decided that while *sensitizer* would be considered an n+s characteristic, it would only be included in the natural language ITD of **Nd:Cr-GSGG laser**.

7.1.9 Part V: Testing the Draft ITDs (Kinds of Laser and Kinds of Neodymium Laser)

With such simple statements, asking the *tout le défini* questions seemed a little excessive, and I did not subject the expert to this process for these two levels. It seemed acceptable to ask the question "Would you feel comfortable using these ITDs to explain these concepts?" Dr. Beckett agreed that each of the three second-level ITDs defined the concepts perfectly. He said that, while the sixth-level ITDs made sense to him, he thought it would be a good idea to contact one more expert (whom he recommended) to verify their validity. I decided to first write up the natural language ITDs and then consult Dr. Claude Rolland. Again, the subgraph and CCMs helped me explain the work that had been done. I then presented him with a list of natural language ITDs, found in Table 10.

CONCEPT	ITD
laser	A device that produces a strong, highly parallel beam of light of a specific wavelength.
gas laser	A laser with a gaseous active medium.
liquid laser	A laser with a liquid active medium.
solid-state laser	A laser with a solid active medium.
semiconductor diode laser	A solid-state laser whose active medium is a semiconductor.
"doped-insulator laser"	A solid-state laser whose active medium is an impurity-doped insulator.
transition metal ion laser	A doped-insulator laser whose active species is a transition metal ion.
rare earth ion laser	A doped-insulator laser whose active species is a rare earth ion.
erbium laser	A rare earth ion laser whose active species is Er^{3+}.
holmium laser	A rare earth ion laser whose active species is Ho^{3+}.
thulium laser	A rare earth ion laser whose active species is Tm^{3+}.
neodymium laser	A rare earth ion laser whose active species is Nd^{3+}.
Nd:YAG laser	A neodymium laser whose host material is yttrium aluminum garnet (YAG).
Nd:YLF laser	A neodymium laser whose host material is yttrium lithium fluoride (YLF).
Nd:Cr-GSGG laser	A neodymium laser whose host material is gadolinium scandium gallium garnet (GSGG), with Cr^{3+} sensitizing the Nd^{3+}.
Nd:glass laser	A neodymium laser whose host material is glass.

Table 10. Draft ITDs (16 concepts).

Dr. Rolland authenticated the ITDs and I entered them into the TKB. The result for **Nd:YAG laser** is shown in Figure 59.

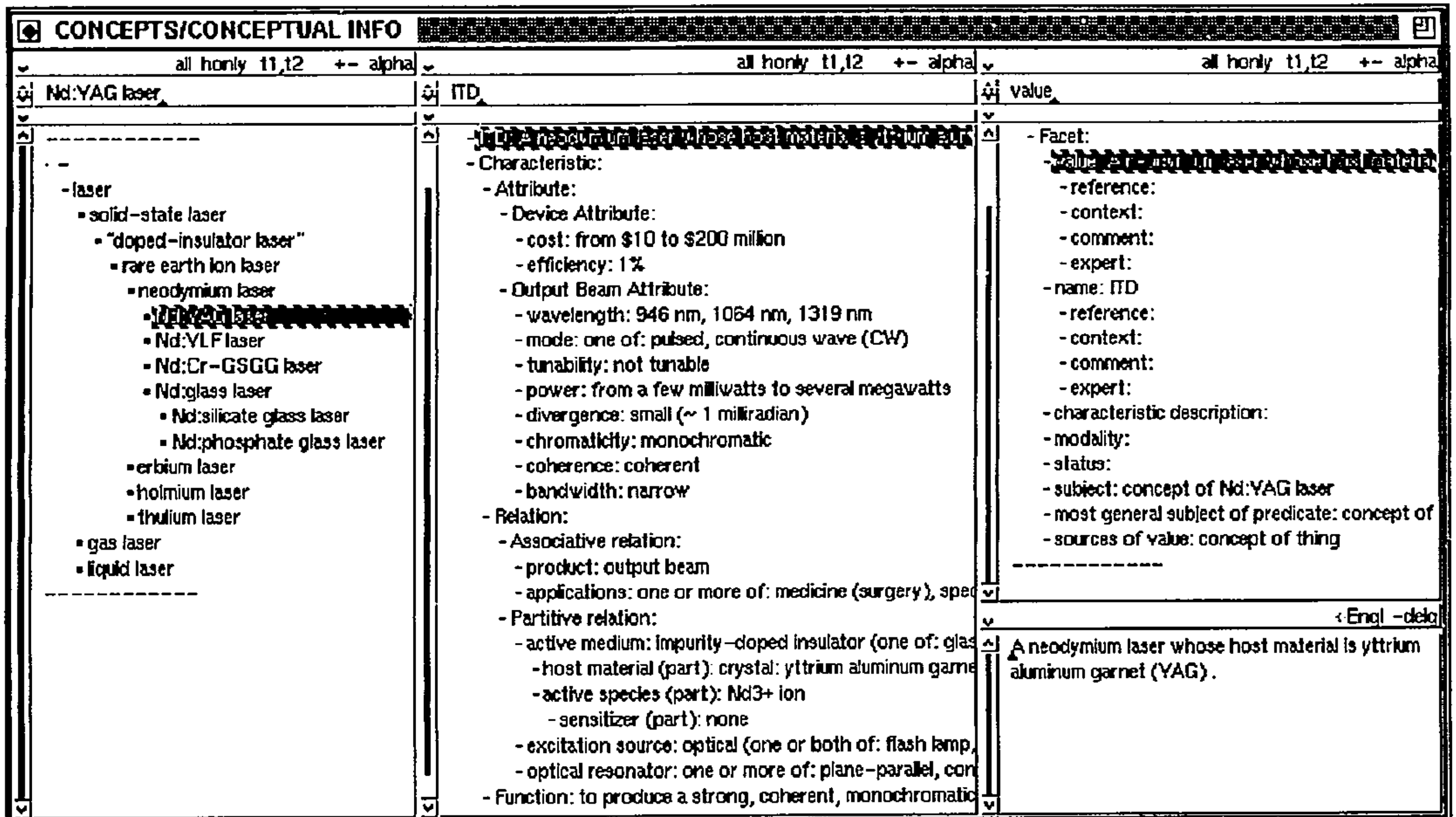


Figure 59. Final ITD for **Nd:YAG laser** in the Concepts/Conceptual Info browser (bottom right pane).

Chapter 8 ASSESSING THE RESULTS: SOME CONCLUDING REMARKS

8.0 INTRODUCTION

This chapter provides an assessment of the results obtained from the test described in chapter 7, including issues that arose, changes and additions to the definition-oriented approach to concept analysis, suggestions for improvements to CODE, and possible directions for future research.

8.1 ISSUES THAT AROSE FROM THE TEST

Two main issues arose from the test: the seeming complexity of it, and the fact that the compound terms used to denote many of the concepts in the subject field were reflected in the final ITD.

8.1.1 The Seeming Complexity of the Approach

When Dr. Beckett was helping me with my definition construction, he often gave the impression of wondering why I was going to so much trouble to come up with such seemingly simple ITDs. It is true that this approach requires that the terminologist perform a fair amount of work, but what is the alternative? As Table 11 shows (and as Table 4 and Table 7 showed in chapter 4), working from contexts, the typical approach in TERMIUM III, results in definitions that are often inconsistent in genus (therefore not providing the user with any indication of the concept system) and inconsistent in differentia (therefore not showing the user how co-concepts differ).

solid-state laser (B)[91]	**gas laser (A)**	**liquid laser (A)**
lasers that use crystals as the active medium (LASRE*1984*20*01*08)	(1) an optical oscillator or amplifier employing as the active medium a gaseous discharge in a suitable gas, vapor, or mixture (AIP-14*1964***13); (2) a laser containing a gaseous lasing medium in a glass tube in which a constant flow of gas replenishes the molecules depleted by the electricity of chemicals used for excitation (COLAS*1980***203) (3) A laser whose active material is gas in a quartz or glass tube with a Brewster angle window at each end. (LASTE*1974***606)	A laser whose active material is dissolved in a liquid contained in a transparent cylindrical shell. (LASTE*1974***606)

Table 11. Definitions for types of **lasers**, extracted from TERMIUM III.

The TERMIUM III definitions in Table 12 are fairly consistent in their genus; however, it is interesting to note that even though these two definitions are both from the same source, they still have inconsistent differentia. The **Nd:YAG laser** ITD includes information about the ***source of excitation*** (both for **Nd:YAG laser** *and* **Nd:glass laser**) while the **Nd:glass laser** ITD discusses ***power***, ***mode***, and ***applications***. On the other hand, the ITDs constructed in chapter 7 (a) reflect the concept system consistently, allowing users to clearly see the relations between the concepts, and (b) present the differentia consistently, allowing users to see what differentiates co-concepts at each level.

Nd: YAG laser (B)	**Nd: glass laser (B)**
A solid-state laser of neodymium: yttrium-aluminum-garnet, similar to the Nd: Glass laser. Both are pumped by flashlamp. (COLAS*1980***206)	A solid-state laser of neodymium: glass offering high power or short pulses, or both, for specific industrial applications. (COLAS*1980***206)

Table 12. Definitions for types of **neodymium lasers**, extracted from TERMIUM III.

[91] TERMIUM files can be one of three different types: A = single concept file (highest quality); B = main file (medium quality); C = working file (unconfirmed quality).

8.1.2 Compound Terms Reflecting N+S Characteristics

It is interesting to note that the compound terms used to label the concepts in the **laser** subject field tend to reflect the n+s characteristics used in the ITDs. Although this may appear to simplify definition construction (i.e. simply use the term denoting the n+s characteristic in the ITD), two points must be made.

First, the compound term does not spell out the n+s characteristic perfectly, even in this particular subject field where compound terms describe their concepts quite clearly. For example, you cannot simply define a **solid-state laser** as *a laser with a solid state.* If the concept is to be understood, the ITD must specify that it is the ***active medium*** which is *solid*, so that a **solid-state laser** is *a laser with a solid active medium.* Similarly, a **Nd:glass laser** is not simply *a laser made of neodymium and glass*; it is *a neodymium laser whose **host material** is glass.* By simply placing the lexical components of the compound term into a ITD, the terminologist risks including the incorrect genus (as was the case with the **Nd:glass laser** example) and misrepresenting the differentia (as occurred with the **solid-state laser** example).

The second point to be made is that, although I found a subject field in which the compound terms somewhat reflect the n+s characteristics of the concepts they denote, not all compound terms work this way. Some compounds obscure the meaning of the concept; Sager [1990:77] provides the example of the term **French drain**. Transforming this compound term into a definition would result in the rather unhelpful *a drain that is French.* Perhaps French drains were originally differentiated from their co-concepts this way, but it's unlikely that this definition would be used today. With a compound term such as **machine translation,** the characteristic may lack precision: one cannot say that **machine translation** is simply *a kind of translation done by a machine.* There are many different machines in the world, but not all of them (e.g. washing machines, sewing machines) are qualified to translate.

Modifiers in compound terms cannot always be trusted to indicate a concept's n+s characteristic(s); I simply happened to be lucky with my chosen subject field (and even then, did not realize this until after all the work was completed and the ITDs constructed).

8.2 CHANGES AND ADDITIONS TO THE APPROACH

While the approach outlined in chapter 6 was successful as a whole, there are some changes and additions that could be made.

8.2.1 Changes to the Approach

The modifications to the approach are the following:

(1) In Part I (Carrying out Preparatory Work), the two steps (d) *if subject field is large, focus on one particular subfield (indicated by "concept cluster")* and (e) *consult experts about (i) preliminary concept graph and (ii) possible documentation*, should be combined. If it is feasible, the terminologist should let the expert help her decide what subfield to start with. Whether the subfield is chosen because it is the expert's area of specialization, or because the expert feels it is fundamental to understanding the rest of the subject field, the expert will likely influence the terminologist's decision. For example, since I had no particular subfield preference, my expert (Dr. Fortin) suggested that I start with **solid-state lasers**, since that was his area of expertise.

(2) In Part III (Carrying Out Concentrated Knowledge Acquisition & Revision), the terminologist should work on groups of co-concepts when revising the collected information and having it verified by experts. As Landau [1984:252] points out, "The only way to avoid contradictory definitions is to draw up entries for a group of related concepts as a group."

This can be difficult at the start of a project, as the documentation does not always have the concept system as clearly divided as the terminologist might like. For example, few documents provided a clear idea of the concept system, and that concepts were grouped together according to no particular pattern. PA91 was the one document that was very consistent; it provided me with a great deal of insight into the various groups of co-concepts by the use of, for example, clear section headings. When a document does not seem to present any sort of concept system, however, it may be better to simply ignore the way it groups concepts in order to avoid confusion. If she is aware that a concept is one of several co-concepts, the terminologist is more likely to start noting which characteristics differentiate the concepts, despite the documentation.

(3) In Part V (Testing the Draft ITD), the "tout le défini" questions seemed rather excessive for the three kinds of **laser** concepts and four kinds of **neodymium laser** concepts. They were useful, however, for the concept of **laser** itself, when there was some doubt about whether ***power*** or ***divergence*** should be considered n+s characteristics. Thus, while this step should remain in the approach, I would note that it does not have to be applied when the information provided by documents and experts is clear.

What the test showed, more than anything, is that a consistently applied approach worked at all levels of the hierarchy. Although the **solid-state laser** subfield appears to get more complicated as it gets more specific, it is really only a few necessary characteristics which dictate the division of each level (and in the case of the various **laser** concepts, it was always the characteristic denoted within the term). I also found that as I moved down through the hierarchy, I became more used to that way of working and the process got easier.

8.2.2 Additions to the Approach

What should be added to the approach is an introduction with important hints for the terminologist. These include the following:

(1) Be sensitive to indications of the characteristics' importance. If the documentation doesn't provide these indications (or if the terminologist suspects that the indications are misleading[92]), the terminologist should consult the expert before entering large amounts of information. It is important that the terminologist not let the technology "run away with her." While CODE makes it fairly easy to record information about a concept, it is necessary to control the number of characteristics recorded. Too much information is almost as bad as too little: it ends up overwhelming the expert (who has to verify the information), the terminologist (who has to decide on the characteristics' modality and create draft definitions), and the end-user, who will eventually have to wade through it all.

Another crucial indication of importance, as this test clearly showed, is the indication of an n+s characteristic in the term. Except for **laser**, all the concepts in this test had their n+s characteristic somehow reflected in their term. This is worth noting for future applications of this approach, especially in subject fields where concepts are labelled with compound terms.

(2) Be aware of your unidimensional restriction. Because multidimensionality is so complicated, this approach instructs the terminologist to divide up the subject field according to the most common dimension. However, the terminologist should be aware of two things. First, subject-field experts are used to looking at the field from a multidimensional perspective (as I noted from my experts' comments when they first saw the unidimensional graph). Second, by consulting a range of documentation, the terminologist is going to be exposed to a variety of ways of dividing up the subject field. Thus, she must ensure that the experts are aware of the unidimensional aspect of her approach (and remind them from time to time, if necessary). She must also take care to keep this restriction in mind herself when scanning documentation. While authors might not feel compelled to represent the

[92] One author (HECHT92), for example, felt that ***applications*** was very important, discussing it for every concept; I thus added it to my list of characteristics. However, in retrospect, I should have paid more attention to the clues in the documentation: HECHT92 was a guidebook about *commercial* **lasers**. For commercial purposes, ***applications*** would be very important. For other purposes, such as simply describing the concept, it would not.

subject field according to one dimension, the terminologist must.

(3) Value your experts. The terminologist should keep in mind how important experts are to this process. An ITD can be described as the tip of an iceberg, showing only a small part of a large amount of information. Only someone who is familiar with every part of the iceberg knows how much should be exposed to fresh air. Ten Pas (1991:78) notes that defining terms is

> a task that cannot and should not be performed by the terminologist alone, but requires the involvement of one or more other experts as well. It is the expert who guarantees the correctness of the **content**, whereas the terminologist is responsible for the **form** of the definition.

(4) Think of yourself as a "knowledge analyst." Vallerga [1991:2] notes that

> [l]e rôle de l'analyste de la connaissance...consiste justement à établir le pont entre l'expertise humaine, avec tout ce qu'elle implique de flou, d'insaisissable et d'indéfinissable, et le données que l'on en veut obtenir et qui seront successivement introduites dan le système à élaborer. L'analyste travaille donc d'une part en relation avec l'expert, dont il observe et analyse le comportement et, d'autre part, en relation avec l'application finale, souvent, ou presque toujours, par l'intermédiaire d'une ordinateur, avec lequel il communique à travers des langages spécifiques plus ou moins évolués.

As point (3) implies, it is the expert who provides the expert knowledge; however, the terminologist must ensure that this knowledge is clearly presented in the TKB.[93] This means that she must carefully guide the expert through the recorded information, including the generic–specific relations and characteristic names, values, and modalities, then elicit the necessary changes from the expert. It is up to the terminologist to carry out quality control and ensure that the TKB presents the subject–field information, including the ITDs, clearly and effectively.

(5) Remember that this approach is meant to be a guide, not a step–by–step recipe. Definition construction is part science and part art. Hence, this approach is intended to guide the terminologist, not force her to carry out all activities in a fixed order. It may be that some steps can

[93] At one point, when I asked Dr. Beckett his opinion of the phrasing of a particular definition, he replied, "Well, it's just English. I mean, that's your speciality, isn't it?" Unwittingly (and quite succinctly), he pointed out one of the most important functions of the terminologist.

be omitted for a particular project or that some can be combined. As this test showed, concept analysis is an iterative process, and many steps will likely have to be repeated throughout a project.

(6) Do not be put off by the seeming complexity of the approach. If readers are given the impression that this approach implies a great deal of work, part of this impression may be caused by the fact that it has been described in great detail. It may seem like a lot of trouble to go to, but there are times when such efforts are necessary. The terminologist who uses this approach will likely construct ITDs that clearly represent the concept system and provide the user with an adequate understanding of the concepts within that system.

The two main points to keep in mind are the following:

(1) Part of the impression of quantity of work is an illusion. The technology helps that terminologist carry out this approach quite quickly, in fact: changes are easy to do, updating between various views is automatic, and the information can be presented in a variety of ways. It is certainly faster than working with pencil-and-paper, and the results are online and shareable. Finally, as I realized during the execution of this test, the more you practice, the better and faster you become.

(2) Nevertheless, this approach does require some work and is based on a more careful, explicit recording of conceptual characteristics than is done in a lot of terminology environments. Thus, it is suitable for those environments where such attention to detail is essential. For example:

— the creation of standards by national and international standards organizations. Standardizing committees often spend hours in meetings, and take months, even years, to create published standards.[94] An approach such as this one could make their job easier.

— the creation of in-house standards or technical definitions by companies. Many companies have commercial products which require a high degree of standardization in the accompanying documentation. For example, BNR, where I currently work, uses the same terminology across a large

[94] Ingrid Meyer, personal communication.

number of software products. From my experience there, I can see that there is a need for standardized terms and good definitions to help developers, technical writers, translators, and clients communicate about the product with each other and avoid what Skuce/Meyer [1990c:132] term the "pass-my-confusion-onto-the-next-person" phenomenon.

It is difficult to say whether working this way really takes more time than the traditional do-it-in-my head or pencil-and-paper approaches: no one has ever "measured," for example, how long it takes to do one term entry and then *correct* the entry later because people have criticized its quality.

7) Take advantage of the technology. The terminologist should know what other computerized tools can be used in conjunction with CODE to facilitate such an approach. These have been summarized in Appendix C.

8.3 SUGGESTIONS FOR IMPROVEMENTS TO CODE

While CODE proved to be enormously helpful in carrying out this definition-oriented approach to concept analysis, there are a few potential areas for improvement, outlined below.

8.3.1 Full Browser Functionality for the CCM

The CCM *must* have the same functionality as the other browsers. Specifically:

(a) the CCM should be resizeable and removable when attached to another browser. At the moment, for example, if a CCM is driven off a graphical browser (which is useful for showing a set of co-concepts and their accompanying differentiating characteristics), the graph and the CCM take up an equal amount of screen space (see Figure 60), and it is impossible to alter their size. Ideally, the CCM should take up the majority of the space, since it presents the majority of the information. It should also be possible to remove the CCM once it is created (leaving only the subgraph).

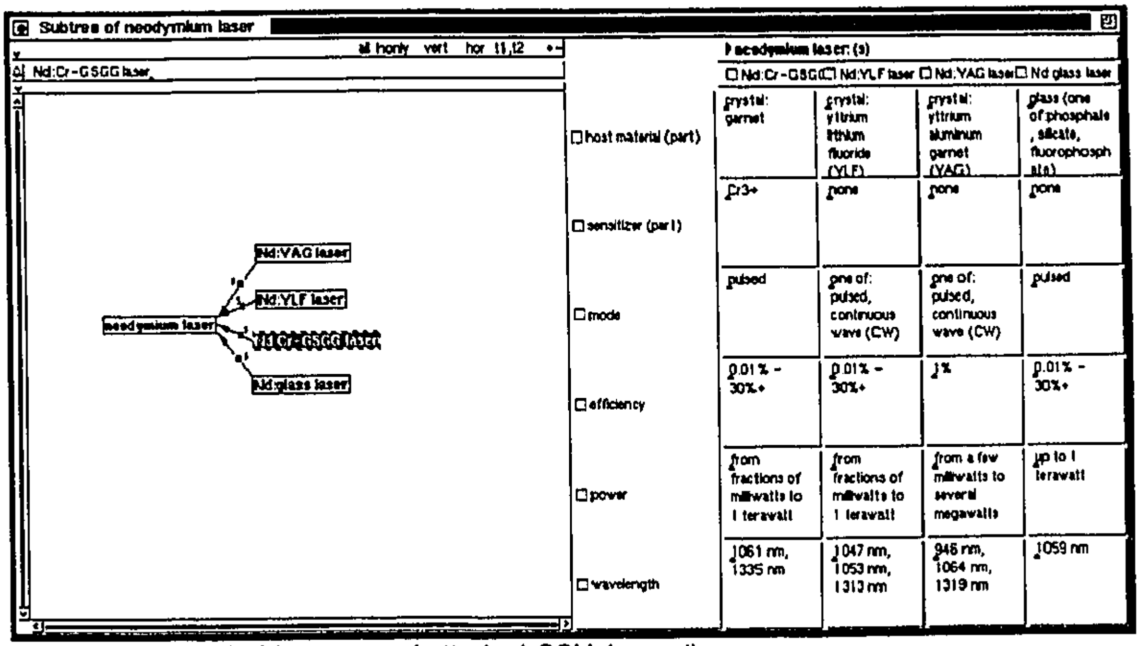


Figure 60. Graphical browser and attached CCM (current).

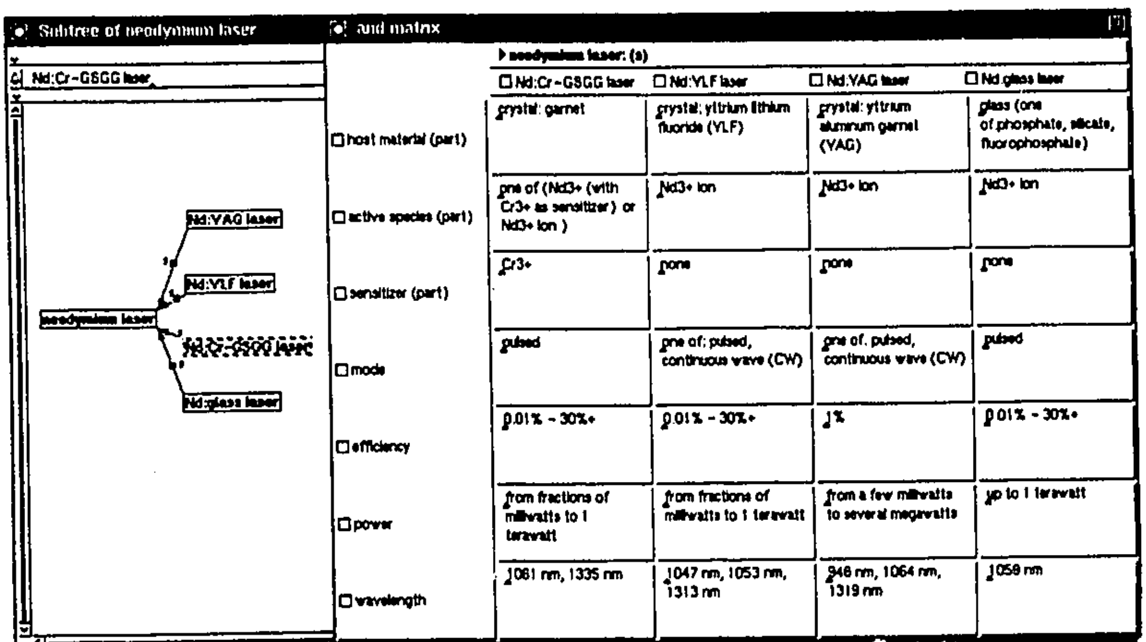


Figure 61. Graphical browser and attached CCM (proposed).

(b) the CCM should be able to have hotkeys (two-key combinations used as shortcuts for specific commands) set for it. This would make some tasks, such as moving characteristic names around, much easier.

(c) the CCM must be able to have facet browsers generated off it so that facets describing a concept's characteristics can be viewed at the same time. This would be *extremely* helpful during knowledge acquisition and verification; otherwise, the terminologist is forced to use the Conceptual Info browser a great deal, because that is where the facets can be seen.

(d) the characteristics in the CCM should appear in a hierarchy, just as they do in the Conceptual Info browser. It is difficult to tell what relation a given characteristic has with other characteristics (e.g. ***active species*** and ***host*** were both parts of ***active medium***, but it was not possible to tell this in the CCM).

8.3.2 Hypertext Links Between the ITD and TKB browsers

It would be very useful to have hypertext links between concept names and characteristic names in the ITD and the other browsers in which they are displayed. This would facilitate browsing between the ITD and the other information in the TKB, which is especially important given that the ITD is supposed to act as a pointer to the rest of the information in the TKB. It would be helpful if the user could simply click on a characteristic in the ITD, e.g. ***active medium*** for **laser**, and be shown that characteristic in the Conceptual Info browser. The user could then look at the attached facets which show, for example, the *VALUE*, *CHARACTERISTIC DESCRIPTION*, and *REFERENCE* of the ***active medium***.

8.3.3 Warning of Modified N+S Characteristics

It would also be quite useful to have some kind of link between the n+s characteristics in the ITD and those in the Conceptual Info browser and the CCM, so that if a characteristic is modified in one of these browsers, the terminologist is signalled to change the ITD as well.

8.4 POSSIBLE DIRECTIONS FOR FUTURE RESEARCH

While this approach makes it much easier for the terminologist to create consistent ITDs in the subject field on which it was tested, there are other areas of research that could be explored to see how the approach could be enhanced. These are discussed below.

8.4.1 Syntax of Definitions

The question of how to transform a set of n+s characteristics into a natural language definition arose during the creation of the ITD for **laser**. The issue of syntax in definitions created in this way is one that warrants more attention.

8.4.2 Multidimensionality

Multidimensionality cries out to be dealt with. Both Dr. Fortin and Dr. Beckett immediately noticed the unidimensionality of the concept system that I showed to them, and took pains to explain to me that, in reality, the neat lines connecting a given concept to only one generic concept should be replaced by a network of criss-crossing lines. This would mean representing the true multidimensionality of the field, something that was beyond the scope of my approach.[95] In the **laser** subject field, however, it is important to realize that **lasers** are not described not only in terms of their ***active medium*** (and its parts), but also in terms of

— ***tunability***: **tunable lasers** can produce beams with a range of wavelengths, **non-tunable lasers** produce beams with only one wavelength

— ***output mode***: **pulsed lasers** produce bursts of ER, while **continuous wave lasers** produce one continuous beam of ER

[95] To the best of my knowledge, no research has ever been done on integrating multidimensionality with the construction of ITDs. Even a superficial familiarity with this problem makes it clear that is highly complex and would likely be too difficult for an M.A. thesis. It could conceivably be the subject of a Ph.D. thesis, however.

— *source of excitation*: e.g. **chemical lasers** are **lasers** whose ***excitation source*** is *chemical*

Ideally, all these dimensions should be represented in the concept system of a **laser** TKB and somehow included in the ITDs. How to go about doing this would make an interesting, albeit difficult, thesis on its own.

8.4.3 Characteristics

More research should be put into the nature of characteristics, especially the different kinds (attributes and relations) and the different modalities (typical, necessary, n+s). It would be useful to have some rules of thumb for recognizing characteristics that are less likely to be necessary (e.g. characteristics with a range of values, such as ***wavelength*** for many of the top **laser** concepts, or characteristics with more than one value, such as ***source of excitation*** for many concepts).

8.4.4 Definition Construction for Other Subject Fields

It would be interesting to see how this particular approach fares in other subject fields, such as those within so-called "soft sciences." These subject fields would likely bring up different and more complex problems: for example, n+s characteristics would probably not be denoted so neatly in compound terms.

8.4.5 Construction of Other Types of Definition

This approach dealt only with the intensional terminological definition; however, there are other types of definition which could be explored. For example, part-whole relations are quite common in terminology. It would be interesting to see if this approach could be adapted to accommodate partitive definitions.

8.4.6 Multilingual Definition Construction

Working on a unilingual project brought up a host of interesting issues, and working with more than one language would certainly create many more. For example, what does one do when the conceptual structures do not match one-to-one? They happened to match in the subject field under study, perhaps because it is within the hard sciences where English predominates as the language of communication. This is not the case in all subject fields, however (e.g. when the same technology develops in parallel in two language groups, such as the military weapons developed during the Cold War).

8.4.7 Guidelines for Solicitation of Expert Help

Experts are a terminologist's most important resource, and it is important that their knowledge and time not be wasted. There is a need to develop an approach for "getting the most out of the expert" which could advise the terminologist on how to go about selecting experts, conducting interviews, etc. This could make a very interesting thesis, especially if it were combined with another specific approach such as the one developed here.

8.4.8 Concordancing Tools

Concordancing tools (e.g. TACT, WordCruncher) allow the terminologist to scan large corpora of online information and produce contexts of specified terms. It would be worthwhile looking into such a system for the scanning process (and even for the preliminary reading process, to verify potentially important concepts or characteristics).[96] With a characteristics template, the terminologist would find such a system especially helpful, as long as it was possible to find the appropriate

[96] Some preliminary work has already been done in this direction within the COGNITERM framework, e.g. Meyer/ Mackintosh in preparation, Meyer/McHaffie 1993.

specialized information in electronic form.

8.5 CONCLUSION

The discussion in this chapter has shown that the proposed definition-oriented approach to concept analysis works well (with some modifications) within a "hard science" subject field such as **lasers**, which has a clear hierarchical structure and explicit characteristics describing its concepts. CODE has proven to be a useful tool for carrying out computerized concept analysis with a large degree of ease and flexibility, allowing the terminologist to represent a clear, consistent concept system; record explicit, consistent characteristics; and create consistent intensional terminological definitions that clearly reflect the system and characteristics.

Aristotle's remarks on the definition still hold true today: it *is* the easiest thing to criticize, appearing so deceptively simple on the surface when well-constructed. Having analyzed ITDs as a product, a process, and tool; developed an definition-oriented approach to concept analysis using CODE; and applied this approach to a practical situation, I now fully appreciate why people have been studying definitions for over 2500 years and yet still not come to a consensus about what they are or how to construct them.

Appendix A CHARACTERISTIC COMPARISON MATRIX:[97] IDEAL VERSION

1. General

The Characteristic Comparison Matrix (CCM) can be opened for any concept selected on the browser. It is a table, and would look a bit like a spreadsheet, as illustrated below:

CHARACTERISTIC COMPARISON MATRIX:
subconcepts of 'furniture for seating people'

→ 1–4 is standard, but some couches can seat up to 6

	(living room furniture) furniture for seating people COUCH	**chair**	**stool**	**pouf**	**bench**
no. persons	1–4 is standard	1	1	1	1–4
have back	yes	yes	no	no	yes or no
padded	yes	yes or no	yes or no	yes	yes or no
conversion to bed	sometimes	n/a	n/a	n/a	n/a
reclinable	n/a	sometimes	n/a	n/a	n/a
swivels	n/a	sometimes	sometimes	n/a	n/a

- **couch** is the concept that was selected in the browser (it should be capitalized or somehow emphasized)

- the user has said she wants to see only those co-concepts that are children of the parent **furniture for seating people**; note that **couch** has another parent as well (**living-room furniture**), which is in brackets above **furniture for seating people** (the co-concepts by this parent are ignored in the display)

- **chair, stool, pouf,** and **bench** are the co-concepts of **couch** (i.e. the concepts that share the same super)

- **no. persons, have back,** etc. are the union of the characteristics of **couch** and its co-concepts *except for* those characteristics that have identical bodies (for this reason, the characteristic *solid:yes* is not here, since it has been inherited from above and is identical in all the co-concepts)

[97] The Characteristic Comparison Matrix (CCM) was initially called the Property Comparison View (with 'property' being used as a synonym for 'characteristic').

- **contents of cells** (a cell is one element of this matrix):
  — if the characteristic exists and has a value, the value is entered in the cell (e.g. ***no. persons*** has a value for all the co-concepts) - if the whole value does not fit in the cell, one can click on the cell and see all of it at the arrow above (as in Excel)
  — if the characteristic does not exist for the concept, the cell indicates n/a (for example, the characteristic ***conversion to bed*** exists only for **couch**, and not for any of the co-concepts)
  — if the characteristic exists, but has no value, the cell is left blank
  — if the characteristic has been blocked for the concept, the cell should indicate BLOCKED (in capital letters)

- ideally, the view should be scrollable

2. What the program should do

- find all the superconcepts [and dimensions] of the concept one is interested in

- if there is more than one, ask the user which one she is interested in (this one is printed directly above the concept name—the ones the user is not interested in are printed directly above that, but in brackets, as on example)

- find all the co-concepts for the concept one is interested in (the concept that is selected on the browser). From this set, **REMOVE:**
  — any concepts whose names start with "related-to" or "other" (they are dummy concepts that should be totally ignored in the CCM)
  — any concepts that are in a different dimension-relation to the super (for example, if your concept is in a dimension2-relation to the selected super, eliminate all concepts that are in a dimension1- or dimension3-relation - only retain the others that are in a dimension2-relation - see **optical film** example, below)

- extract the union of all conceptual information characteristics of the concept and its co-concepts (i.e. only the characteristics that belong to the category conceptual information - ignore all characteristics in any other category)

- from this union, eliminate all the characteristics which have identical bodies in the co-concepts

- WARNING: if the concept has a characteristic, but the co-concepts DO NOT, you still want to retain that characteristic (and mark n/a in the appropriate cells); similarly, if one or more of the concept's co-concepts have a characteristic that the concept does not, you still want to retain that characteristic (and mark n/a in the appropriate cell)

NOTE:
The only characteristics that should *not* show up are the ones where the bodies are identical.

- there should be an option available in the CCM that lets one view these identical characteristics if one wishes. A separate window should open up, entitled "Characteristics Identical in Co-concepts." This window should list the characteristics and bodies in question, without mentioning the concepts.

Example:

Characteristics Identical in Co-concepts

solid:yes
weight:a weight
purpose:for sitting
location:inside a building

3. A real example of what we want, taken from the optical storage TKB
NOTE: the concept-to-be-defined is **optical film.**

CHARACTERISTIC COMPARISON MATRIX
subconcepts of 'optical storage media'

→ common, flexible photographic film

	(read-only media) optical storage media **OPTICAL FILM**	**optical tape**	**optical disc**	**optical card**
content	video images and compressed audio	one or more of: data, audio, graphics, still pictures and motion video	one or more of: data, audio, graphics, still pictures and motion video	one or more of: data, audio, graphics, still pictures and motion video
encoding method	digital and/or analog	digital and/or analog	digital and/or analog	digital
material	common, flexible photographic film			
physical form	film	tape	disc	card
storage capacity	up to 30,000 frames in NTSC	1 TB (12-inch reel)		from 2 MBs to 16 MBs (an equivalent of 800 pp of text)
dimensions/ diameter	n/a	n/a	3.5, 4.72, 5.25, 8, 12, or 14 inches	n/a
dimensions/ area	n/a	n/a	n/a	size of a credit card

- **optical film** has 2 supers: **optical storage media**, and **read-only media**; here the user has indicated she is interested in the super **optical storage media**, so **read-only media** is in brackets

- co-concepts by super **optical storage media**: **optical tape, optical disc, digital paper, optical card** (dimension2)

  NOTE: **read-only media, WORM media, erasable media** (dimension1) are ignored, since **optical film** is in dimension2

- co-concepts by super **read-only media**: **optical tape, optical card-read only, ROM disc**

  NOTE: **optical-card read only** and **ROM disc** are ignored, since the user has specified she wants the co-concepts for the super **optical storage media**

4. Tim Lethbridge's work estimates for Ken Iisaka to do the Ideal Characteristic Browser

	V2 (hrs)	Abbreviated version (hrs)
Newview		
integration into existing	2	2
simple textual listing	1	1
cellular layout	8	
cell selectable	4	
scrolling if cell selectable	3	
identical props pop-up	3	
API		
basic algorithm	3	2
identical props	2	
	26 hrs	6 hrs
overhead 50%	13 hrs	3 hrs
TOTALS	39 hrs	9 hrs

5. QUESTIONS on Characteristic Comparison Matrix (to be thought about as it is used)

- would it be helpful to have the superconcept included on the table?

- if a characteristic is introduced in the concept-to-be-defined, but does not exist anywhere else, is it at all relevant to definition construction? (example: ***conversion to a bed*** is introduced for a **couch**, but does not exist for **chair, stool, pouf, bench** - it doesn't seem very relevant to the definition, but perhaps in another concept, a similar phenomenon might be relevant - can we get pertinent examples?)

- if this kind of characteristic is never important for the definition, then should the CCM not include it at all?

- if a characteristic exists in one or more co-concepts, but not in the concept-to-be-defined, is it relevant for the definition? Example: ***reclines*** is introduced in **chair**, but nowhere else; ***swivels*** is introduced in both **stool** and **chair**, but nowhere else. An extreme case of this phenomenon is where the characteristic exists in all (or at least the majority of) co-concepts, but is blocked in the concept-to-be-defined, for example, for **penguin** — that it ***flies*** - in this case, the blocked characteristic is very important

- it seems as though the most critical characteristics are the ones that exist in all co-concepts, but where the bodies are somewhat different in them - the BIG QUESTION is, from this group of characteristics, how to decide which are the most important ones to put in the definition?

6. The following is a summary of all the inheritance possibilities that can be seen so far, and how they appear on the Characteristic Comparison Matrix:

CASE 1:

characteristic and value inherits identically to all subconcepts - TOTALLY irrelevant for definition - does not normally show on CCM unless one specifically asks for it

CASE 2:

characteristic inherits but values NOT the same in co-concepts - this is probably the most important case for definitions, and shows on CCM

CASE 3:

characteristic introduced in concept-to-be-defined, but not in co-concepts (or at least, not in very many of them) Example: ***conversion to bed***, above - it doesn't seem terribly important, but we need more examples to analyze (these characteristics are indicated on CCM)

CASE 4:

characteristic introduced in one or more co-concepts, but not in concept-to-be-defined. Example ***reclines***, ***swivels***, above. Seems that this is not terribly important for definition, but we need more examples (these characteristics are indicated on CCM)

CASE 5:

characteristic blocked in concept-to-be-defined: indicated on CCM, and probably *very* important for definitions

Appendix B THE LASER TKB

The following is an output of the laser TKB. Figure 62 shows the concepts that were examined in detail for the test of the definition-oriented approach to concept analysis.

CONCEPT NAME: LASER

Conceptual Information:

ITD: A device that produces a coherent, monochromatic beam of electromagnetic radiation.
commercial value: from $10 to $200 million
reference: LAUR86:102, TF87:18
comment: suggested by FORTIN93
modality: t

wavelength: [x-ray to microwave]
reference: LUPA92:90, LAUR86:6
comment: value should be in nanometres (nm) where documented
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)
modality: n

mode: one or both of: pulsed, continuous wave (CW)
reference: HECHT92:487
context: Table: Commercial Laser Types, Organized by Wavelength
expert: FORTIN93
reference: MESS90:84
context: Table: Characteristics of Nd-based lasers
comment: HECHT92:487 refers to this as "output type" (within Table: Commercial Laser Types, Organized by Wavelength)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)
modality: n

tunability: one of: tunable, not tunable
reference: TF87:16
context: description of dye lasers, and the advantage they have over other lasers
comment: FORTIN93 comments that most lasers are not tunable; the exceptions are dye lasers and solid-state tunable lasers.
expert: FORTIN93
reference: TF87:16, KOE91:66
comment: DULEY91:139 defines tunable as "having an output wavelength that can be adjusted over a range of 1 nm or more."

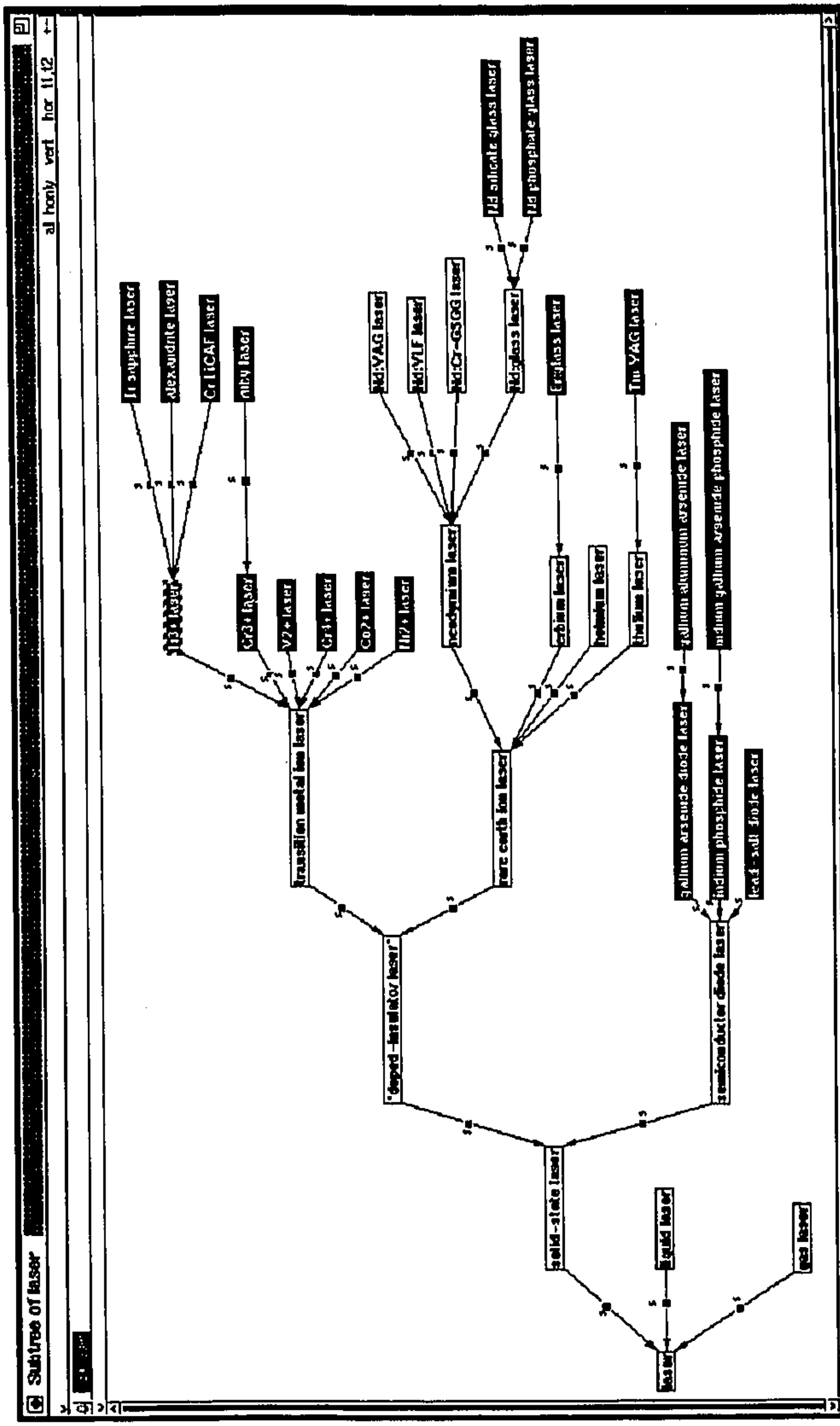


Figure 62. Graph from the laser TKB. Concepts examined in the test are denoted by white nodes with black letters.

expert: FORTIN93
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
reference: TF87:9
comment: no light source, the laser included, is capable of producing absolutely monochromatic light (OSHEA77:12); LUPA92:91 explains that high monochromaticity means there is small wavelength spread
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.
modality: n+s

coherence: coherent
reference: HECHT92:4
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)
modality: n+s

divergence: low
reference: OSHEA77:13
comment: OSHEA77 : 13 notes that this value is for a small laser; BECKETT93 says that some lasers (e.g. semiconductor lasers) have a fairly high divergence
expert: BECKETT93
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)
modality: t

power: from fractions of milliwatts to megawatts
reference: HECHT92:2
comment: Power is often found in ranges because the value depends on whether the laser is cw or pulsed (i.e. whether the laser is in a long pulse, Q-switched or mode locked). (adapted from TF87:30); measured in watts (W) (TF87:2)
expert: FORTIN93
reference: TF87:2, LAUR86:6
context: Table : Laser Properties and Applications (TF87 : 30)
expert: FORTIN93
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)
modality: t

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation

emitted.
modality: n

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.
reference: LAUR86:20,102
comment: literally thousands of laser applications have evolved from scientific research (LAUR86:20)
modality: t

active medium: one of: gas, liquid, solid
reference: FORTIN93
context: interview with expert
expert: FORTIN93
reference: HECHT86:341, TF87:10, PA91:603
context: glossary item, description, definition
comment: synonyms: lasing medium (LAUR86:3), laser medium (HECHT92:3)
expert: FORTIN93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

excitation source: one of: electrical, chemical, nuclear, optical (flash lamp, another laser)
reference: LAUR86:3
context: found within definition and in diagram (LAUR86:3-4)
expert: BECKETT93
reference: TF87:30
context: Table: Some Common Lasers and their Properties; other possible names are 'excitation device,' 'means of excitation' and 'pumping device' (LAUR86:3)
comment: synonyms: power supply (HECHT92:3), means of excitation (LAUR86:3), pumping device (LAUR86:3), source of energy (TF87:10)
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable), none
reference: LAUR86:50
context: Figure: Common types of optical resonators
comment: syn:optical cavity (HECHT92:3); equivalent to the tuner used to tune a radio (FORTIN93); BECKETT93 notes that not all lasers have to have an optical resonator (e.g N2 laser) as the active medium gets excited enough to produce an output beam without one.
expert: FORTIN93

reference: LAUR86:49
context: described as one of the laser's three key elements, along with active medium and excitation device (HECHT92:3)
comment: syn:optical cavity (HECHT92:3); equivalent to the tuner used to tune a radio (FORTIN93)
expert: FORTIN93
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: t

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER
reference: SILF91:209
modality: t

Linguistic Information:

English term: laser
reference: LAUR86:3
context: definition (see def1)

syn1: Light Amplification by Stimulated Emission of Radiation
reference: TF87:4

syn2: laser device
reference: LAUR86:66, HECHT92:2, SIEG86:2

syn3: laser system
reference: PA91:603

def1: A laser is a device that uses a means of excitation to pump energy into the electrons of a lasing medium, which in turn emits and amplifies a light beam formed by mirrors at each end of the lasing medium.
reference: LAUR86:3

def2: In the simplest terms, a laser can be described as a source of light or 'radiant energy'. It has, however, many special properties which make it quite distinct from other light sources such as the sun, a candle, an electric light bulb or a fluorescent tube.
reference: TF87:2

def3: A device that emits light rays that are in phase, travelling in the same direction and essentially of the same wavelength (i.e. colour). A laser beam does not diverge by a significant amount and maintains a high energy density. A laser resides in the CD-ROM reader unit, and is at the heart of all optical technology.
reference: MITCH91:518

def4: Lasers are devices that generate or amplify coherent radiation at frequencies in the infrared, visible, or ultraviolet regions of the electromagnetic spectrum.

reference: SIEG86:2

def5: ...a laser can be considered as a source of a narrow beam of monochromatic, coherent light in the visible, infrared or ultraviolet parts of the spectrum.
reference: HECHT92:2

def6: The laser is basically a light source. The radiation that it emits is not fundamentally different than any form of electromagnetic radiation. The nature of the device, however, is such that some remarkable properties of light are realized.
reference: LUPA92:90

def7: A laser is a device that amplifies, or increases, the intensity of light, producing a strong, highly directional or parallel beam of light of a specific wavelength.
reference: SILF91:209

French term: laser
reference: EIST72:282
gender: m

Background Information:

subconcepts: (concept of solid-state laser, concept of gas laser, concept of liquid laser)

superconcepts: concept of manufactured SER
expert: FORTIN93

classification status: classified

introduction date: 1960 (solid-state ruby laser)
reference: PA91:603

CONCEPT NAME: SOLID-STATE LASER

Conceptual Information:

ITD: A laser with a solid active medium.

commercial value: from $10 to $200 million
modality: t

wavelength: [x-ray to microwave]
expert: BECKETT93
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)
modality: n

mode: one or both of: pulsed, continuous wave (CW)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)
modality: n

tunability: one of: tunable, not tunable
expert: BECKETT93
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.
modality: n

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)
modality: n

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)
modality: t

power: from fractions of milliwatts to 1 terawatt
reference: ECK
comment: range is taken from specific concept values
expert: BECKETT93
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)
modality: n

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.
modality: n

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy),

optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.
modality: t

active medium: solid (one of:impurity-doped insulator, semiconductor)
reference: PA91:603, 604
context: defining context
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n+s

excitation source: one of: electrical, optical (flash lamp, another laser)
reference: ECK93
comment: based on values of specific concepts
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
reference: BECKETT93
expert: BECKETT93
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: t

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER
modality: n

Linguistic Information:

English term: solid-state laser
reference: LAUR86:54

syn1: solid state laser
reference: TF/87:12

syn2: solid laser
reference: LUPA92:149

def1: A solid-state laser is a device in which the active medium is based on a solid material.
reference: PA91:603

def2: A solid-state laser is one in which the active medium is a nonconductive crystal, a crystalline

material, or glass doped with a species that can emit laser light.
reference: HECHT92:96

French term: laser à solide
reference: GRA92:30, EIST72:287
gender: m

syn1: laser à cristaux
reference: JOHN82:16

syn2: laser à plasma solide
reference: JOHN82:17

syn3: laser à milieu solide
reference: GRA92:9

syn4: laser solide
reference: BTQ93

syn5: laser à corps solide
reference: NENT86:174, BTQ93

Background Information:

subconcepts: (concept of "doped-insulator laser", concept of semiconductor diode laser)

superconcepts: concept of laser
expert: FORTIN93

classification status: classified

introduction date: 1960 (ruby laser)
reference: PA91:603

CONCEPT NAME: "DOPED-INSULATOR LASER"

Conceptual Information:

ITD: A solid-state laser whose active medium is an impurity-doped insulator.

commercial value: from $10 to $200 million

wavelength: [x-ray to microwave]
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
expert: FORTIN93
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.
coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZJL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of milliwatts to 1 terawatt
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one or both of: glass, crystal)
reference: PA91:604, OSHEA77:148
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)

modality: n+s

host material (part): one or both of: glass, crystal
reference: OSHEA77:148, 151
comment: the host material is a crystalline insulator _in most cases_ (OSHEA77:148)
expert: BECKETT93
name: host material (part)
reference: LAUR86:16
comment: syn:host medium (LUPA92:92), host (PA91:604). While LUPA92:93 says that some laser types have no host (e.g. the semiconductor diode laser and ion lasers like argon and krypton), BECKETT93 feels that only doped-insulator lasers should have the active medium divided up into host material and active species. characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n

active species: laser-active impurity ion
reference: PA91:604
context: description of this type of laser
comment: syn:impurity, dopant (OSHEA77:148); it is the impurity ions that are responsible for the laser action (PA91:604)
expert: BECKETT93
name: active species
reference: LAUR86:16
comment: syn:laser species (PA91:603), lasing medium (LUPA90:92), lasant (LUPA90:92)
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n

excitation source: optical (one or both of: flash lamp, another laser)
reference: ECK93
comment: based on values of specific concepts
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: "doped-insulator laser"

reference: OSHEA77:148
comment: BECKETT93 says that, although he has never seen this term, he does feel that the concept is useful for dividing up the subfield; thus it has been kept as a "dummy concept"
expert: BECKETT93 says that, although he has never seen this term, he does feel that the concept is useful for dividing up the subfield; thus it has been kept as a "dummy concept"

def1: The term 'doped-insulator laser' is used to describe a laser whose active medium is a regular array of atoms: a crystal with impurity ions intentionally introduced into the crystal at the time of growth (a process called _doping_).
reference: OSHEA77:148

def2: The term 'solid-state laser' is also used for these lasers. However, because the semiconductor laser is a solid-state device, we prefer to label the lasers under discussion as doped-insulator lasers. This term should reduce confusion and serve as a more descriptive label.
reference: OSHEA77:148

Background Information:

subconcepts: (concept of transition metal ion laser, concept of rare earth ion laser)

superconcepts: concept of solid-state laser
reference: OSHEA77:148
comment: OSHEA77:148 says although this kind of laser is also called a solid-state laser, this term is used to avoid confusion; FORTIN93 concurs
expert: FORTIN93, BECKETT93

classification status: classified

CONCEPT NAME: TRANSITION METAL ION LASER

Conceptual Information:

ITD: A doped-insulator laser whose active species is a transition metal ion.

commercial value: from $10 to $200 million

wavelength: 660 nm - 2300 nm
reference: ECK
comment: range is taken from specific concept values
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can

operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of milliwatts to 1 terawatt
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one or both of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): one or both of: glass, crystal
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make

the lasing possible. (adapted from PA91:603)
modality: n

active species: transition metal ion
reference: PA91:610
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n+s

excitation source: optical (one or both of: flash lamp, another laser)
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: transition metal ion laser
reference: PA91:910

Background Information:

subconcepts: (concept of Ti3+ laser, concept of Cr3+ laser, concept of V2+ laser, concept of Cr4+ laser, concept of Co2+ laser, concept of Ni2+ laser)

superconcepts: concept of "doped-insulator laser"
reference: PA91:910
context: subsection title
expert: BECKETT93

classification status: classified

CONCEPT NAME: RARE EARTH ION LASER

Conceptual Information:

ITD: A doped-insulator laser whose active species is a rare earth ion.

commercial value: from $10 to $200 million

wavelength: 160 nm - 4800 nm
reference: PA91:609
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of milliwatts to 1 terawatt
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g.

battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one or both of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): one or both of: glass, crystal
reference: PA91:608
context: descriptions of various kinds of rare earth ion lasers
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n

active species: rare earth ion
reference: PA91:607
context: description of number of RE ions that have been lased
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n+s

excitation source: optical (one or both of: flash lamp, another laser)
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: rare earth ion laser
reference: PA91:607

syn1: RE laser
reference: PA91:607

Background Information:

subconcepts: (concept of neodymium laser, concept of erbium laser, concept of holmium laser, concept of thulium laser)

superconcepts: concept of "doped-insulator laser"
reference: OSHEA77:151
comment: The neodymium laser is said to be a very important kind of doped-insulator laser; since neodymium laser is a kind of RE laser, then it seems logical that an RE laser is a kind of doped-insulator laser.
expert: FORTIN93

classification status: classified

CONCEPT NAME: NEODYMIUM LASER

Conceptual Information:

ITD: A rare earth ion laser whose active species is Nd3+.

commercial value: from $10 to $200 million

wavelength: 1047 nm - 1064 nm [infrared]
reference: see context note
context: range of values is derived from values of specific concepts, [OSHEA77:151]
comment: almost all neodymium lasers operate at 1.064 um (LAUR86:83)
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of milliwatts to 1 terawatt
reference: MESS90:82, 88
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one or both of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): one or both of: glass, crystal
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n

active species: Nd3+ ion
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n+s

sensitizer (part): one or none of: Cr3+
reference: HECHT92:391

expert: BECKETT93
modality: t

excitation source: optical (one or both of: flash lamp, another laser)
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: neodymium laser
reference: LAUR86:80, HECHT92:389

syn1: Nd3+ laser
reference: PA91:607

def1: The neodymium laser consists of neodymium ions (ions are simply atoms with a few electrons missing) embedded in a host material. There are many hosts that have been used for neodymium but the most common are yttrium aluminum garnet (YAG) and glass. The names of the lasers are abbreviated as Nd:YAG and Nd:glass.
reference: LAURENCE86:81

def2: Nd3+ lasers are by far the most technologically important type of lasers. The three main materials that are routinely utilized with Nd3+ are Y3Al5012 and LiYF4 (known as Nd:YAG and Nd:YLF, respectively) and several kinds of glass.

French term: laser à neodyme
reference: HART74:56

syn1: laser au néodymium
reference: GRIFF83:8

Background Information:

subconcepts: (concept of Nd:YAG laser, concept of Nd:YLF laser, concept of Nd:Cr–GSGG laser, concept of Nd:glass laser)

superconcepts: concept of rare earth ion laser
reference: PA91:607, KOE91:36

comment: MESS90:81 says it is a kind of solid-state laser
expert: FORTIN93

classification status: classified

CONCEPT NAME: ND:YAG LASER

Conceptual Information:

ITD: A neodymium laser whose host material is yttrium aluminum garnet (YAG).

commercial value: from $10 to $200 million

wavelength: 946 nm, 1064 nm, 1319 nm
reference: MESS90:84
context: Table: Characteristics of Nd-based Lasers
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
reference: LAUR86:81
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from a few milliwatts to several megawatts
reference: SILFVAST91:221
comment: TF87:30 gives a value of 'up to 1 terawatt; 100 W average c.w.' (Table: Laser Properties and Applications); LFNOV88:50 cites 400 W as the maximum output power; LAUR86:82 cites a few milliwatts to 100+ watts (c.w.) and up to 5,000 megawatts (pulsed)
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): crystal: yttrium aluminum garnet (YAG)
reference: LAUR86:81
context: description of three main types of material utilized with Nd3+
comment: syn:Y3Al5012 (PA91:608)
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n+s

active species: Nd3+ ion
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n

sensitizer (part): none
expert: BECKETT93
modality: n

excitation source: optical (one or both of: flash lamp, another laser)
expert: BECKETT93

characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: Nd:YAG laser
reference: LAURENCE86:47
context: "common solid-state laser"

syn1: neodymium in yttrium aluminum garnet laser
reference: LAURENCE86:81

def1: The most popular neodymium laser is Nd:YAG. ... The use of the YAG crystal, with its thermal stability, permits continuous operation.
reference: LAURENCE86:81

French term: laser YAG
reference: EIST72:287, BTQ93

syn1: laser néodyne-YAG
reference: GRA92:14

syn2: laser au grenat d'yttrium et d'aluminium
reference: BTQ93

syn3: laser yttrium-aluminium-grenat
reference: BTQ93

Background Information:

superconcepts: concept of neodymium laser
reference: MESS90:81
expert: FORTIN93

classification status: classified

CONCEPT NAME: ND:YLF LASER

Conceptual Information:

ITD: A neodymium laser whose host material is yttrium lithium fluoride (YLF).

commercial value: from $10 to $200 million

wavelength: 1047 nm, 1053 nm, 1313 nm
reference: PA91:613 (first two values), MESS90:84 (third value)
context: Table: Properties of Some Common Laser Materials, Table: Characteristics of Nd-based Lasers
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
reference: MESS90:84
context: Table: Characteristics of Nd-based Lasers
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of milliwatts to 1 terawatt
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): crystal: yttrium lithium fluoride (YLF)
reference: MESS90:84
context: description of three main types of material utilized with Nd3+
comment: syn: LiYF4 (PA91:608)
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n+s

active species: Nd3+ ion
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n

sensitizer (part): none
expert: BECKETT93
modality: n

excitation source: optical (one or both of: flash lamp, another laser)
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: Nd:YLF laser
reference: HECHT92:390

syn1: Nd:yttrium lithium fluoride laser
reference: LFJUNE90:84

Background Information:

superconcepts: concept of neodymium laser
reference: MESS90:81
expert: FORTIN93

classification status: classified

CONCEPT NAME: ND:CR-GSGG LASER

Conceptual Information:

ITD: A neodymium laser whose host material is (GSGG), with Cr^{3+} sensitizing the Nd^{3+}.

commercial value: from $10 to $200 million

wavelength: 1061 nm, 1335 nm
reference: PA91:613, MESS90:84
context: Table: Properties of Some Common Laser Materials, Table: Characteristics of Nd-based Lasers
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
reference: LFJUNE90:84
context: Table: Characteristics of Nd-based Lasers
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)

modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of milliwatts to 1 terawatt
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): crystal: gadolinium scandium gallium garnet
reference: PA91:613
context: Table: Properties of Some Common Laser Materials
comment: syn: GSGG (PA91:613), Gd3Sc2Ga3012 (PA91:613)
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n+s

active species: Nd^{3+} ion
reference: PA91:613
context: Table: Properties of Some Common Laser Materials
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr^{3+} ions in a ruby laser. (LUPA92:92)
modality: n

sensitizer (part): Cr^{3+}
reference: HECHT92:391
context: description of concept
expert: BECKETT93
modality: n+s

excitation source: optical (one or both of: flash lamp, another laser)
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: Nd:Cr–GSGG laser
reference: HECHT92:390

syn1: Cr,Nd:GSGG laser
reference: PA91:613, LFJUNE90:84

syn2: Nd:Cr:GSGG laser
reference: LFNOV88:31

syn3: Nd,Cr:GSGG
reference: MESS90:84

Background Information:

superconcepts: concept of neodymium laser
reference: MESS90:81
expert: FORTIN93

classification status: classified

CONCEPT NAME: ND:GLASS LASER

Conceptual Information:

ITD: A neodymium laser whose host material is glass.

commercial value: from $10 to $200 million

wavelength: 1059 nm
reference: MESS90:84
context: Table: Characteristics of Nd-based Lasers
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
reference: MESS90:84
context: Table: Characteristics of Nd-based Lasers
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: up to 1 terawatt
reference: MESS90:88

characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): glass (one of:phosphate, silicate, fluorophosphate)
reference: HECHT86:81
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n+s

active species: Nd3+ ion
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n

sensitizer (part): none
expert: BECKETT93
modality: n

excitation source: optical (one or both of: flash lamp, another laser)
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified

by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: Nd:glass laser
reference: LAUR86:81, MESS90:84

def1: The largest laser in the world is a neodymium in glass (Nd:glass) laser.
reference: LAUR86:81

syn1: Nd-glass laser
reference: HECHT92:390

syn2: neodymium in glass laser
reference: LAUR86:81

French term: laser à verre au neodyne
reference: EIST72:285

Background Information:

subconcepts: (concept of Nd:silicate glass laser, concept of Nd:phosphate glass laser)

superconcepts: concept of neodymium laser
reference: MESS90:81

classification status: classified

CONCEPT NAME: ERBIUM LASER

Conceptual Information:

ITD: A rare earth ion laser whose active species is Er3+.

commercial value: from $10 to $200 million

wavelength: 1610 nm [infrared]
reference: TF87:13, PA91:614
context: Table: Active Element and Laser Wavelength
comment: PA91:609 says that the range is 0.18 to 5.2 um (although many of these are not commercial systems)
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of milliwatts to 1 terawatt
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): one or both of: glass, crystal
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n

active species: Er3+ ion
reference: PA91:605
context: Table: Laser Ions and Abbreviated Listing of Host Materials
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n+s

excitation source: optical (one or both of: flash lamp, another laser)
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: erbium laser
reference: KOE91:37

syn1: Er3+ laser
reference: PA91:608
context: title of subsection within Rare Earth Ion Laser section

French term: laser erbium
reference: TERMIUM93
comment: B file

Background Information:

subconcepts: concept of Er:glass laser

superconcepts: concept of rare earth ion laser
reference: PA91:608, KOE91:37
expert: FORTIN93

classification status: classified

CONCEPT NAME: HOLMIUM LASER

Conceptual Information:

ITD: A rare earth ion laser whose active species is Ho3+.

commercial value: from $10 to $200 million

wavelength: 2050 nm
reference: TF87:13
context: Table: Active Element and Laser Wavelength
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.

coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of milliwatts to 1 terawatt
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): one or both of: glass, crystal
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n

active species: Ho3+ ion
reference: PA91:605
context: Table: Laser Ions and Abbreviated Listing of Host Materials
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n+s

excitation source: optical (one or both of: flash lamp, another laser)
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: holmium
reference: HECHT92:471

French term: laser holmium
reference: TERMIUM93
comment: B file

Background Information:

superconcepts: concept of rare earth ion laser
reference: KOE91:37
expert: FORTIN93

classification status: classified

CONCEPT NAME: THULIUM LASER

Conceptual Information:

ITD: A rare earth ion laser whose active species is Tm3+.

commercial value: from $10 to $200 million

wavelength: 1120 nm
reference: TF87:13
context: Table: Active Element and Laser Wavelength
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic

characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.
coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of milliwatts to 1 terawatt
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.

active medium: impurity-doped insulator (one of: glass, crystal)
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n

host material (part): one or both of: glass, crystal
expert: BECKETT93
characteristic description: The material which is host to the ion whose energy transitions make the lasing possible. (adapted from PA91:603)
modality: n

active species: Tm3+ ion
reference: PA91:605
context: Table: Laser Ions and Abbreviated Listing of Host Materials
expert: BECKETT93
characteristic description: The atoms or molecules that actually emit the light, such as the Cr3+ ions in a ruby laser. (LUPA92:92)
modality: n+s

excitation source: optical (one or both of: flash lamp, another laser)
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: thulium laser
reference: HECHT92:471?

Background Information:

subconcepts: concept of Tm:YAG laser

superconcepts: concept of rare earth ion laser
reference: KOE91:38
expert: FORTIN93

classification status: classified

CONCEPT NAME: SEMICONDUCTOR DIODE LASER

Conceptual Information:

ITD: A solid-state laser whose active medium is a semiconductor.

commercial value: from $10 to $200 million

wavelength: 330 nm - 400 nm
reference: LUPA92:150
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

mode: one or both of: pulsed, continuous wave (CW)
reference: TF87:30
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of

light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: not tunable
expert: FORTIN93
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.
coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)

power: from fractions of microwatts to 100 watts
reference: LAUR86:95, FORTIN93
comment: Although LAUR86:95 had the top limit as 1000 watts, FORTIN93 feels that this figure is too high, and that the top limit is probably about 100 watts.
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)

product: output beam
reference: LAUR86:94
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: optical storage, communications technology, military (e.g. battlefield simulation systems, weapons)
reference: LAUR86:94

active medium: semiconductor
reference: HECHT92:97
comment: The number of semiconductor crystals was specified by FORTIN93.
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341–2)
modality: n+s

excitation source: electrical
reference: TF87:17
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: semiconductor diode laser
reference: LAURENCE86:61
context: listed under heading "Lasers that don't work the way we just said they do"; described as the most common type of laser in use, very simple and very inexpensive, and therefore very important

syn1: semiconductor laser
reference: LAURENCE86:19

syn2: diode laser
reference: LAUR86:19

syn3: semi-conductor laser
reference: TAYLOR/FRENCH87:17
context: listed as one of the two other most important kinds of laser (i.e. semi-conductor and chemical) beside solid-state, gas and dye

def1: The term solid-state laser is not meant to include semiconductor lasers which, even though solid, operate on entirely different principles
reference: LAUR86:81

def2: A solid-state laser is one in which the active medium is a nonconductive crystal, a crystalline material, or glass doped with a species that can emit laser light. Semiconductor lasers are considered functionally different types, even though they are made of solid-state materials, because of fundamental differences in their operation.
reference: HECHT92:96

def3: The application of semiconductor diode lasers (themselves solid-state lasers, but normally classified separately because of the different pumping mechanism of current injection) for pumping other solid-state laser materials ... may have been the most important development in the process to create compact, reliable lasers....

reference: LFJUNE90:90

French term: laser à semi-conducteur
reference: GRA92:30, NENT86:168

syn1: laser semi-conducteur
reference: JOHN82:16, GRIFF83:8

syn2: laser à diode
reference: JOHN82:16

Background Information:

subconcepts: (concept of gallium arsenide diode laser, concept of indium phosphide laser, concept of lead-salt diode laser)

superconcepts: concept of solid-state laser
reference: TF87:30
context: Table: Laser Properties and Applications
comment: many authors distinguish semiconductor diode lasers from solid-state lasers; however, FORTIN93 says they definitely are a kind of solid-state laser
expert: FORTIN93, BECKETT93

classification status: classified

introduction date: 1962
reference: LAUR86:95
comment: after the semiconductor diode laser was invented, ten years of reascarch and development were required to produce ones that could be operated reliability at room temperature

CONCEPT NAME: GAS LASER

Conceptual Information:

ITD: A laser with a gaseous active medium.

commercial value: from $10 to $200 million

wavelength: ˜ 100 nm – ˜ 2000 nm [x-ray to far infra-red]
reference: TF87:30, OSHEA77:126
context: Table: Laser Properties and Applications
comment: value is taken from a range of values; DULEY91:139 comments that output wavelengths span the spectrum from the far infrared to the
expert: BECKETT93
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)

modality: n

mode: one or both of: pulsed, continuous wave (CW)
reference: TF87:30
context: Table: Laser Properties and Applications
comment: value is taken from a range of values
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: one of: tunable, not tunable
reference: DULEY91:139
comment: DULEY91:139 comments that few gas lasers are tunable over an appreciable wavelength range
expert: BECKETT93
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.
coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)
modality: t

power: up to 10 megawatts
reference: TF87:30
context: Table: Laser Properties and Applications
expert: BECKETT93
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)
modality: t

product: output beam
context: general description
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation

emitted.
modality: n

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.
modality: t

active medium: gas
reference: TF87:2
context: explanation of three main types of laser
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341–2)
modality: n+s

excitation source: one of: electrical, chemical, optical (flash lamp, another laser)
reference: DULEY91:139
context: description of how gas lasers are pumped (introductory paragraph)
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable), none
expert: BECKETT93
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: t

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER
modality: n

Linguistic Information:

English term: gas laser

French term: laser à gaz
reference: MCKIE84:15, HART74:59

syn1: laser gaz
reference: BTQ93

Background Information:

subconcepts: (concept of excimer laser, concept of metal-vapor laser, concept of helium-neon laser, concept of carbon dioxide laser, concept of noble gas ion laser, concept of free-electron laser)

superconcepts: concept of laser
expert: FORTIN93

classification status: classified

CONCEPT NAME: LIQUID LASER

Conceptual Information:

ITD: A laser with a liquid active medium.

commercial value: from $10 to $200 million

wavelength: 200 nm - 1800 nm [gamma ray to infra-red]
expert: BECKETT93
characteristic description: The distance between adjacent crests and troughs of a wave of radiation. (adapted from LAUR86:23)
modality: n

mode: one or both of: pulsed, continuous wave (CW)
characteristic description: Output characteristic of laser; normally refers to the intensity distribution of output light intensity. (DULEY91:139) Lasers can emit single pulses (i.e. one pulse of light each time they are fired); they can emit trains of pulses, one short pulse after another; or they can operate continuously: the terms 'pulsed' (also 'single pulse operation') and 'continuous wave' (also 'continuous wave operation') are used to describe these conditions. (LAUR86:29)

tunability: tunable
reference: TF87:16
expert: FORTIN93
characteristic description: Ability of lasers to change their wavelength over a wide range. (TF87:16)
modality: n

chromaticity: monochromatic
characteristic description: Whether the beam is made up of one or more wavelengths; monochromatic means one wavelength, while polychromatic means more than one wavelength.
coherence: coherent
characteristic description: The relative organization of the light waves in the light signal (coherent means the separate light waves in the beam are exactly in step with one another, i.e. they have the same phase; incoherent means they aren't in step). The term can apply to any form of

electromagnetic radiation. (adapted from LAUR86:32, ZIL73:5)

divergence: low
characteristic description: The spreading out of a beam wtih distance, measured as an angle. (adapted from HECHT86:347)
modality: t

power: from fractions of milliwatts to megawatts
expert: BECKETT93
characteristic description: The rate at which the light source radiates energy; measured in watts (W). (TF87:2)
modality: n

product: output beam
characteristic description: The type of electromagnetic radiation emitted from the source, i.e. visible radiation, invisible radiation (gamma rays, x-rays, ultraviolet light, infrared radiation, microwaves, radio waves). Related to wavelength in that the wavelength dictates the type of radiation emitted.

applications: one or more of: medicine (surgery), spectroscopy (the study of the behaviour of atoms and molecules by exciting them with light and observing how they absorb or reemit the light energy), optical storage, industry (welding, drilling), electronics (trim, anneal and microsolder the microcircuits in computers, televisions, radios, etc.), energy production (experimental: e.g. initiation of fusion reactions), holography, optical communications technology, non-linear optics, research, military (e.g. battlefield simulation systems, weapons), interferometry, etc.
modality: t

active medium: liquid
reference: TF87:2
context: explanation of three main types of laser
expert: BECKETT93
characteristic description: The substance that originates and amplifies the laser light. Generally, the name of the laser identifies the active medium. (adapted from LAUR86:3, HECHT86:341-2)
modality: n+s

excitation source: optical (flash lamp, another laser)
reference: LAUR86:87
expert: BECKETT93
characteristic description: The laser's source of energy, which helps it achieve population inversion.
modality: n

optical resonator: one or more of: plane-parallel, confocal, hemispherical (may be stable, unstable)
reference: BECKETT93
characteristic description: The set of two mirrors at each end of the lasing medium; their function is to feed back the spontaneously emitted photons which can then oscillate and be amplified by stimulated emission to form the laser beam.
modality: n

Function: to produce a coherent, monochromatic (i.e. of one specific wavelength) beam of ER

Linguistic Information:

English term: dye laser

French term: laser à liquide
reference: NENT86:101

syn1: laser liquide
reference: GRIFF83:8, BTQ93

syn2: laser à plasma liquide
reference: JOHN82:17

Background Information:

subconcepts: concept of dye laser

superconcepts: concept of laser
expert: FORTIN93

classification status: classified

Appendix C TECHNOLOGY TO BE USED WITH CODE

It was noted in chapter 8 that the terminologist should explore the various tools that can be used in conjunction with CODE in order to facilitate computerized concept analysis. Some of these tools include:

— **wordprocessing applications such as WP5.1 (on desktop or notebook computers)**: these applications let the terminologist take notes at libraries or at home during preliminary reading, transfer them to a UNIX workstation, and then cut and paste pertinent information into CODE.[1]

— **text editors:** these UNIX-based applications allow the terminologist to take notes while working on CODE, and then print them or transfer them to a PC. For example, I used a text editor called Xedit to compile lists of questions for the experts as I used CODE. I also made notes while working on the TKB and then transferred them to a PC for inclusion in this thesis. CODE will also convert all the TKB information to a text file. The file can then be transferred to a wordprocessing application to be edited and presented to the expert. That is how the TKB records in Appendix B were created.

— **workspaces:** these empty windows can be opened in CODE for use as electronic notepads, where one can type in important but temporary information, such as reminders about things to do with the TKB (e.g. verify a certain characteristic value).

— **screendumps:** these are files created by "capturing" the windows that appear onscreen. They can either be sent straight to the printer (which is useful for gathering information to be presented to the expert), or placed in a directory. From there they can be transferred (after conversion) to, for example, a personal computer, where they can be imported into a wordprocessing application. This is how the CODE graphics in this thesis were created.

[1] CODE runs on workstations that use the UNIX operating system.

Bibliographic References

Terminology and Knowledge Engineering

Auger, Pierre and Louis-Jean Rousseau. 1978. *Méthodologie de la recherche terminologique.* Québec: Office de la langue française.

Blampain, Daniel, Philippe Petrussa, Marc Van Campenhoudt. 1992. "À la recherche d'écosystèmes terminologiques." *L'environnment traductionnel : La station de travail du traducteur de l'an 2001.* Journées scientifiques du Réseau thématique de recherche «Lexicologie, terminologie, traduction.» Mons 25-27 avril 1991. Eds. André Clas and Hayssam Safar. Sillery, Québec: Presses de l'Université du Québec. 273-282.

Boutayeb, S. 1991. *Catégories de la terminologie et terminologie de l'informatique documentaire.* Mémoire en vue de l'obtention du diplôme d'études approfondies, sciences du langage. Paris: Université de la Sorbonne Nouvelle, Paris III.

Boutin-Quesnel, Rachel, Nycole Bélanger, Nada Kerpan and Louis-Jean Rousseau (Comité consultatif canadien du CT 37 de l'ISO). 1978. *Vocabulaire systématique de la terminologie.* Montréal: Office de la langue française, Gouvernement du Québec.

Bowker, Lynne. 1992. "An Approach to Dealing with Multidimensionality in a Terminological Knowledge Base." Master's Thesis. School of Translation and Interpretation, University of Ottawa.

Bowker, Lynne and Ingrid Meyer. 1993. "Beyond Textbook Concept Systems: Handling Multidimensionality in a New Generation of Term Banks." Paper to appear in the *Proceedings of the Third International Congress on Terminology and Knowledge Engineering (TKE '93)*, to be held in Cologne, Germany, 25-27 August 1993.

Cole, Wayne D. 1987. "Terminology: Principles and Methods." *Computers and Translation.* Vol. 2. 77-87.

de Bessé, Bruno. 1992. "Des fichiers terminologiques aux bases de connaissances." *L'environnment traductionnel : La station de travail du traducteur de l'an 2001.* Journées scientifiques du Réseau thématique de recherche «Lexicologie, terminologie, traduction.» Mons 25-27 avril 1991. Eds. André Clas and Hayssam Safar. Sillery, Québec: Presses de l'Université du Québec. 283-300.

Dubuc, Robert. 1985. *Manuel pratique de terminologie.* 2e édition, revue et augmentée. Québec: Linguatech.

Eck, Karen. 1992. "The Characteristic Comparison Matrix: Present Applications and Future Possibilities." Unpublished Report. University of Ottawa, School of Translation and Interpretation.

Freibott, Gerhard and Heid, Ulrich. 1990. "Terminological and lexical knowledge for computer-aided translation and technical writing." *Proceedings of the Second International Congress on Terminology and Knowledge Engineering Applications.* Frankfurt: INDEKS Verlag. 522-535.

Galinski, Christian. 1991. "Terminology Standardization." *L'actualité terminologique/ Terminology update.* 24 (3). 4-5.

———. 1988a. "Terminology and Information/Knowledge Management." *Languages at Crossroads* (Proceedings of the 29th Annual Conference of the American Translators Association). Ed. Deanna Lindberg Hammond. Medford, NJ: Learned Information Inc. 485-492.

———. 1988b. "Recent Developments in Knowledge Engineering and the International Division of Labor within TermNet." *Languages at Crossroads* (Proceedings of the 29th Annual Conference of the American Translators Association). Ed. Deanna Lindberg Hammond. Medford, NJ: Learned Information Inc. 493-499.

ISO Technical Committee TC-37 *Terminology (principles and co-ordination).* 1991. "Development and Representation of Concept Systems." International Organization for Standardization (Working Document).

———. 1990a. *International Standard: Terminology-Vocabulary (ISO1087:1990 (E/F).* Geneva: International Organization for Standardization.

———. 1990b. "Concept Systems (Development and Representation)." International Organization for Standardization (Working Document).

Meyer, Ingrid. 1992. "Concept Management for Terminology: A Knowledge Engineering Approach." *Standardizing Terminology for Better Communication: Practice, Applied Theory and Results.* Eds. R.A. Strehlow and S.E. Wright. Philadelphia: American Society for Testing and Materials. 140-151.

———. 1991. "Knowledge Management for Terminology-Intensive Applications: Needs and Tools." *Lexical Semantics and Knowledge Representation. First SIGLEX Workshop Proceedings.* Eds. James Pustejovsky and Sabine Bergler. Berlin: Springer Verlag. 21-37.

Meyer, Ingrid, Lynne Bowker and Karen Eck. 1992a. "COGNITERM: An Experiment in Building A Knowledge-Based Term Bank." *EURALEX '92 Proceedings (Part I).* Eds. Hannu Tommola, Krista Varantola, Tarja Salmi-Tolonen, Jürgen Schopp. Tampere, Finland: Tampereen Yliopisto. 159-172.

———. 1992b. "Constructing a Knowledge-Based Term Bank: Some Issues and Implications." *Proceedings from the International Symposium on Terminology and Documentation in Specialized Communication (T&D 91).* Ottawa: Minister of Supply and Services Canada. 232-256.

Meyer, Ingrid and Kristen Mackintosh. 1993 (in preparation). "Phraseological Analysis and Concept Analysis: Exploring a Symbiotic Relationship." To be submitted to EURALEX '93.

Meyer, Ingrid and Bruce McHaffie. 1993. "From Peepholes to Ballrooms: Exploring New Vistas in Terminological Data Representation." Working Paper No. 93-01 of the Artificial Intelligence Lab, Department of Computer Science, University of Ottawa.

Meyer, Ingrid and Line Paradis. 1991. "Applying Knowledge Engineering Technology to Terminology: A Pilot Project." *L'Actualité terminologique/Terminology Update.* 24 (2). 3-7.

Meyer, Ingrid, Douglas Skuce, Lynne Bowker and Karen Eck. 1992a. "Methodological Issues in Knowledge-Based Term Bank Construction." *Proceedings of the fifteenth International Conference on Computational Linguistics (COLING 92), Nantes, France, 25-29 July 1992.* Ed. Christian Boitet. 4 vols. 956-960.

———. 1992b. *COGNITERM: Structure and Methodology.* Technical Report of the Artificial Intelligence Laboratory, Dept. of Computer Science, University of Ottawa, Canada.

Michaud, Diane. 1991. "La normalisation terminologique au service des organisme de normalisation technique au Canada." *L'actualité terminologique/Terminology update.* 24 (3). Ottawa: Department of the Secretary of State of Canada. 6-7.

Miller, David R. 1993. "Towards Knowledge-Base Systems for Translators." Master's Thesis. School of Translation and Interpretation, University of Ottawa.

Nirenburg, Sergei, Ira Monarch, Todd Kaufmann, Irene Nirenburg, and Jaime Carbonell. 1988. *Acquisition of Very Large Knowledge Bases: Methodology, Tools and Applications.* (Technical Report CMU-CMT-88-108.) Pittsburgh: Carnegie Mellon University.

Pearson, Jennifer. 1993. "Terminology and Extra-linguistic Knowledge." Report 2 of the ET10/66 Consortium, National Centre for Language Technology, Dublin City University, Dublin.

———. 1992. "Terminology and Extra-linguistic Knowledge." (2 Vols.) Report 1 of the ET10/66 Consortium, National Centre for Language Technology, Dublin City University, Dublin.

Picht, Heribert and Jennifer Draskau. 1985. *Terminology: An Introduction.* Guildford, Surrey: University of Surrey.

Rey, Alain. 1992. *Que sais-je? La terminologie : noms et notions.* 2nd ed., revised. Paris: Presses Universitaires de France.

Richard, Guy. 1977. "La normalisation notionnelle et la normalisation terminologique." *Actes du sixième colloque international de terminologie.* Pointe-au-Pic (Québec), du 2 au 7 octobre 1977. Québec: Éditeur officiel du Québec. 81-104.

Rondeau, Guy. 1981. *Introduction à la terminologie.* Montréal: Centre Éducatif et Culturel inc.

Rousseau, Louis-Jean. 1991. "La normalisation terminologique, outil au service de la communication." *L'actualité terminologique/Terminology update*, 24 (3). Ottawa: Department of the Secretary of State of Canada. 3.

Sager, Juan C. 1990. *A Practical Course in Terminology Processing*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Skuce, Douglas. 1992. "CODE version 4.1: A System for Knowledge Management Acquisition, Analysis and Retrieval." Publicity Document. AI Laboratory, Department of Computer Science, University of Ottawa.

———. 1991a. "Managing Software Design Knowledge: A Tool and an Experiment". Submitted to *The IEEE Transactions on Knowledge and Data Engineering*.

———. 1991b. "A Wide Spectrum Knowledge Management System". Submitted to *Knowledge Acquisition*.

Skuce, Douglas and Ingrid Meyer. 1990a. "Computer-Assisted Concept Analysis: An Essential Component of a Terminologist's Workstation". *TKE '90: Terminology and Knowledge Engineering. Proceedings of the Second International Congress on Terminology and Knowledge Engineering*. Eds. Hans Czap and Wolfgang Nedobity. 2 vols. Frankfurt: INDEKS Verlag. 187-199.

———. 1990b. "Concept Analysis and Terminology: A Knowledge-Based Approach to Documentation". *Proceedings of the Thirteenth International Conference on Computational Linguistics (COLING-90)*. Ed. Hans Karlgren. Helsinki, Finland: Helsinki University. 56-58.

———. 1990c. "Computer-Assisted Concept Analysis for Terminology: A Framework for Technological and Methodological Research." *EURALEX '90 Proceedings. IV International Congress*. Malaga, Spain: BIBLIOGRAF. 129-140.

Sowa, John. 1991. "Conceptual Analysis as a Basis for Knowledge Acquisition." *The Cognition of Experts: Psychological Research and Empirical AI*. Ed. R.R. Hoffman. Berlin: Springer Verlag.

Talbot, George and O'Croinin, Donncha. 1992. "Developing EIRETERM: The Termbank of the Eurotra Machine Translation Project. *Proceedings of the Fifth Euralex International Congress*. Tampere, Finland: University of Tampere. 179-184.

TLS (Terminology and Linguistic Services). 1984. *The Terminologist's Handbook* (Draft). Ottawa: Terminology and Linguistic Services, Official Languages and Translation, Department of the Secretary of State of Canada.

Vallerga, Paola. 1991. "La consultation d'expert de domaine. Principes et techniques de la sollicitation des connaissances." Mémoire de D.E.A. Institut national des langues et civilisations orientales. Centre de recherche en ingénierie multilingue (C.R.I.M.).

Winston, Morton E., Roger Chaffin, and Douglas Herrmann. 1987. "A Taxonomy of Part-Whole Relations." *Cognitive Science*, (11), 417–444.

Wüster, Eugen. 1968. *The Machine Tool. An Interlingual Dictionary of Basic Concepts.* London: Technical Press.

DEFINITIONS

Aristotle. 1928. "Topica and De Sophisticis Elenchis." *The Works of Aristotle*. Vol. 1. Trans. W.A. Pickard. Ed. W.D. Ross. First ed. London: Oxford University Press.

Borsodi, Ralph. 1967. *The Definition of Definition: A New Linguistic Approach to the Integration of Knowledge.* Boston, Massachusetts: Porter Sargent Publisher.

Dahlberg, Ingetraut. 1987. "Computerization of Definition Texts: A New Approach." (Paper given at the 7th Int. SCCAC-Meeting, Trier, 28 Sept. 1987. To be published in *Content-Concepts-Meaning. Publications from the Society of Conceptual and Content Analysis by Computer (SCCAC)*, Greenwich, CT: JAI Press, 1990 (Advances in Computing and the Humanities, Vols. 3 and 4)). Ed. K. Schmidt. 16–22.

———. 1983. "Terminological Definitions: Characteristics and Demands." *Problèmes de la définition et de la synonymie en terminologie.* Actes du Colloque international de terminologie. Université Laval, Québec, 23–27 mai 1982. Ed. Diane Duquet-Picard. Québec: GIRSTERM. 15–34.

———. 1981. "Conceptual Definitions for INTERCONCEPT." *International Classification*, 8 (1). 16–22.

de Bessé, Bruno. 1990. "La définition terminologique." *La Définition*. Actes du Colloque *la Définition*, organisé par le CELEX (Centre d'études du lexique) de l'Université Paris-Nord (Paris 13, Villetaneuse) à Paris, les 18 et 19 novembre 1988. Eds. Jacques Chaurand and Francine Mazière. Paris: Larousse. 252-261.

Harris, Brian. 1983. "The Need for Definitions on Term Records: A Translator's View." *Problèmes de la définition et de la synonymie en terminologie*. Actes du Colloque international de terminologie. Université Laval, Québec, 23-27 mai 1982. Ed. Diane Duquet-Picard. Québec: GIRSTERM. 141-164.

Hermans, A. 1989. "La définition des termes scientifiques." *Méta*. (34) 3. 529-532.

Lo Pipero, F. 1990. "Aristote : la syllabe comme modèle de la signification de de la définition." *La Définition*. Actes du Colloque *la Définition*, organisé par le CELEX (Centre d'études du lexique) de l'Université Paris-Nord (Paris 13, Villetaneuse) à Paris, les 18 et 19 novembre 1988. Eds. Jacques Chaurand and Francine Mazière. Paris: Larousse. 24-29.

Mel'čuk, Igor. 1988. "Principes et critères de Description Sémantique dans le DEC." *Dictionnaire explicatif et combinatoire du français contemporain (Recherches Lexico-Sémantique II).* Montréal: Les Presses de l'Université de Montréal. 27-39.

Miller, John William. 1980. *The Definition of the Thing, with Some Notes on Language.* New York: W.W. Norton & Company.

Nkwenti-Azeh, Blaise. 1989. "An investigation into the structure of the terminological information contained in special language definitions." Diss. University of Manchester.

Rey-Debove, J. 1966. "La définition lexicographique : recherches sur l'équation sémique." *Cahiers de Lexicologie.* VIII (1). Paris: Didier-Larousse. 71-94.

Robinson, Richard. 1965. *Definition.* Oxford: Oxford University Press.

Rousseau, Louis-Jean. 1983. "La définition terminologique." *Problèmes de la définition et de la synonymie en terminologie.* Actes du Colloque international de terminologie. Université Laval, Québec, 23-27 mai 1982. Ed. Diane Duquet-Picard. Québec: GIRSTERM. 35-46.

Sager, Juan. 1983. "Definitions in Terminology." *Problèmes de la définition et de la synonymie en terminologie.* Actes du Colloque international de terminologie. Université Laval, Québec, 23-27 mai 1982. Ed. Diane Duquet-Picard. Québec: GIRSTERM. 113-139.

ten Pas, Elsemiek. 1991. "On the role of the Terminologist in the Definition of Terms." *Tools for the Knowledge Organization and the Human Interface* (Proceedings of the 1st International ISKO-Conference, Darmstadt, 14-17 August 1990). Ed. Robert Fugmann. Vol. 2. Frankfurt/ Main: INDEKS Verlag. 78-84.

Lexicology/Lexicography

Atkins, B.T.S. 1991. "Building a Lexicon: The Contribution of Lexicography." *International Journal of Lexicography.* 4 (3).

Béjoint, Henri. 1988. "Scientific and Technical Words in General Dictionaries." *International Journal of Lexicography,* 1 (4). Oxford: Oxford University Press. 354-368.

Boguraev, Branimir K. and Beth Levin. 1990. "Models for Lexical Knowledge Bases." *Electronic Text Research. Proceedings of the Sixth Annual Conference of the UW Centre for the New Oxford English Dictionary and Text Research.* Waterloo, Canada: University of Waterloo. 65-78

Cruse, D.A. 1986. *Lexical semantics.* New York: Cambridge University Press.

Hartmann, R.R.K. 1983. "Theory and Practice in Dictionary-Making." *Lexicography: Principles and Practice.* Ed. R.R.K. Hartmann. London: Academic Press. 3-11.

Landau, Sidney. 1984. *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.

Miller, George A. 1990. "Nouns in WordNet: A Lexical Inheritance System." *International Journal of Lexicography*, 3 (4). 245–264.

Pustejovsky, James and Sabine Bergler (Eds.). 1991. *Lexical Semantics and Knowledge Representation. First SIGLEX Workshop Proceedings.* Berlin: Springer Verlag.

Shaikevitch, A. and I. Oubine. 1988. "Translators and researchers look at bilingual terminological dictionaries." *Babel*, 34 (1). 10–16.

LASERS

ASIMOV90: Asimov, Isaac. 1990. *How Did We Find Out About Lasers?* New York: Walker and Company.

BECKETT93:[1] Beckett, Dr. Douglas. June–July 1993. Personal interviews.

DULEY91: Duley, W.W. 1991. "Gas Lasers." *Encyclopedia of Lasers and Optical Technology.* San Diego: Academic Press, Inc. 139–148.

EIST72: 1972. "Laser." *Encyclopédie internationale des sciences et des techniques*. 10 vols. Paris: Presses de la Cité.

FORTIN93: Fortin, Dr. Emory. May 1993. Personal interviews.

GRA92: Graham, Ian and Phillippe Chandelon. 1992. *Comment fonctionne... Les lasers et les hologrammes*. Tournai: Éditions Gamma.

GRIFF83: Griffiths, John. 1983. Lasers et Hologrammes. Paris: Éditions Garnier Frères.

HECHT86: Hecht, Jeff. 1986. *The Laser Guidebook.* New York: McGraw–Hill, Inc.

HECHT92: Hecht, Jeff. 1992. *The Laser Guidebook.* 2nd ed. New York: McGraw–Hill, Inc.

JOHN82: Johnson, James. 1982. *Le laser*. Paris: Bordas.

JOHN81: Johnson, James. 1981. *A Look Inside Lasers*. Milwaukee: Raintree Publishers Inc.

KOE91: Koechner, Walter. 1991. *Solid–State Laser Engineering.* 3rd completely revised and updated ed. Berlin: Springer–Verlag.

[1] The laser references have been given codes to facilitate the work done in the TKB. These codes are used both in the TKB and in chapters 7 and 8 of the thesis.

LAUR86: Laurence, Clifford L. *The Laser Book. A New Technology of Light.* New York: Prentice Hall Press. 1986.

LFNOV88: *Laser Focus.* November 1988. 24 (11).

LFJUNE90: *Laser Focus.* June 1990. 26 (6).

LFAUG87: *Laser Focus.* August 1987. 23 (8).

LUPA92: Luxon, James T. and David E. Parker. 1992. *Industrial lasers and their applications.* 2nd ed. Englewood Cliffs, NJ: Prentice Hall.

MCKIE84: McKie, Robin and François Carlier. 1984. *Les lasers.* Tournai: Éditions Gamma.

MESS90: Messenger, Heather W. August 1990. "Solid-state lasers develop new capabilities." *Laser Focus.* 26 (8). 81-97.

MITCH91: Mitchell, Joan (Ed.). 1990. *The CD-ROM Directory 1991 (5th Edition).* London: TFPL Publishing.

NENT86: Nentwig, Karl Heinz. 1986. *Elsevier's Dictionary of Opto-Electronics and Electro-Optics.* Amsterdam: Elsevier Science Publisher B.V.

OSHEA77: O'Shea, Donald C., W. Russell Callen and William T. Rhodes. 1977. *Introduction to Lasers and their Applications.* Reading, MA: Addison-Wesley Publishing Company.

PA91: Payne, Stephen A. and George F. Albrecht. 1991. "Solid-State Lasers." *Encyclopedia of Lasers and Optical Technology.* San Diego: Academic Press, Inc. 603-624.

ROLLAND93: Rolland, Dr. Claude. 15 July 1993. Personal interview.

SIEG86: Siegman, Anthony E. 1986. *Lasers.* Mill Valley, California: University Science Books.

SILF91: Silfvast, William T. 1991. "Lasers." *Encyclopedia of Lasers and Optical Technology.* San Diego: Academic Press, Inc. 209-224.

TF87: J.R. Taylor and P.M.W. French. 1987. *How Lasers are Made.* New York: Facts on File Publications.

ZIL73: Zilczer, Paul. 1973. *Laser Technology: A Survey.* Washington, DC: Technology Utilization Office, NASA.

Documentation Used for Examples

Chen, Ching-chih. 1989. *Hypersource on Optical Technologies.* Chicago: American Library Association.

Cringely, Robert X. 1992. *Accidental Empires. How the Boys of Silicon Valley Make Their Millions, Battle Foreign Competition, and Still Can't Get a Date.* Reading, Massachusetts: Addison-Wesley Publishing Company, Inc.