Proceedings of DETC98
1998 ASME Design Engineering Technical Conference
September 13-16, 1998, Atlanta, GA

DETC98/DTM-5665

DESIGN INFORMATION RETRIEVAL: IMPROVING ACCESS TO THE INFORMAL SIDE OF DESIGN

WILLIAM H. WOOD, MARIA C. YANG, MARK R. CUTKOSKY
Center for Design Research
Stanford University

ALICE M. AGOGINO
Department of Mechanical Engineering
University of California, Berkeley

ABSTRACT

Capturing and reusing design experience holds great potential for improving designer effectiveness. The first step toward leveraging lessons from the past for design decision making is gaining access to them. Because decisions early in the design process largely determine its ultimate success, it is important to embrace the informal, unstructured information that is prevalent during conceptual design. Information retrieval is proposed as the basis for access to this informal design information. By creating hierarchical thesauri of life cycle design issues, design process terms, and component and system functional decompositions, we hope to establish an intermediate language in which design context can be captured. Experiments in design information retrieval exploiting design context for determining document similarity within design case studies and design notebooks demonstrate the value of this approach.

1 INTRODUCTION

While effort directed toward formalizing information during the design process has made significant gains, the earliest stages of the design process have steadfastly resisted systematic formalization. While there appears to be no theoretical reason for this, there may be significant social barriers to formalizing conceptual design. Clearly, formalization is an approach with potentially tremendous payoff in terms of automating the design process. However, we believe that significant gains can be made by impressing some level of formalism over design information as it now typically exists and that this effort might also aid in the formalization process.

This two-pronged approach has an analog in computer science/information retrieval. There are efforts underway to formalize the total of human knowledge in the hope of capturing within the computer that which is known so that natural language will become machine understandable, a rather formidable task (Lenat, 1995). With the boom of the World Wide Web and its seemingly endless supply of unstructured or loosely structured documents, improving access to information stored in 'natural' languages has received a great deal of attention on both commercial and academic fronts.

It may be hyperbolic to compare the task of formalizing design knowledge with that of formalizing human knowledge in general. Circumscribing the target domain to design may make formalization tenable; our focus is on how such circumscription can be exploited to improve access to informal design information. We take a lesson from work in an even more highly circumscribed domain: design for manufacture.

Industry's successful application of design for manufacture in limited domains provided the imetus for DFM formalization efforts which, in turn, spawned software capable of evaluating the manufacturability of significant subsets of mechanical design (Beiter et al., 1993, BDI, Hryniak et al., 1996). However, little widespread adoption in industry has resulted from this formalization. Instead, industry has largely implemented a team-based approach to design which emphasizes informal information sharing among design team stakeholders. It seems that formalization of mechanincal engineering faces similar obstacles to those which have stagnated general-purpose artifical intelligence. This is not to say that such efforts will not eventually prove valuable, simply that perhpas embracing the informal aspects of design might be a valid alternative.

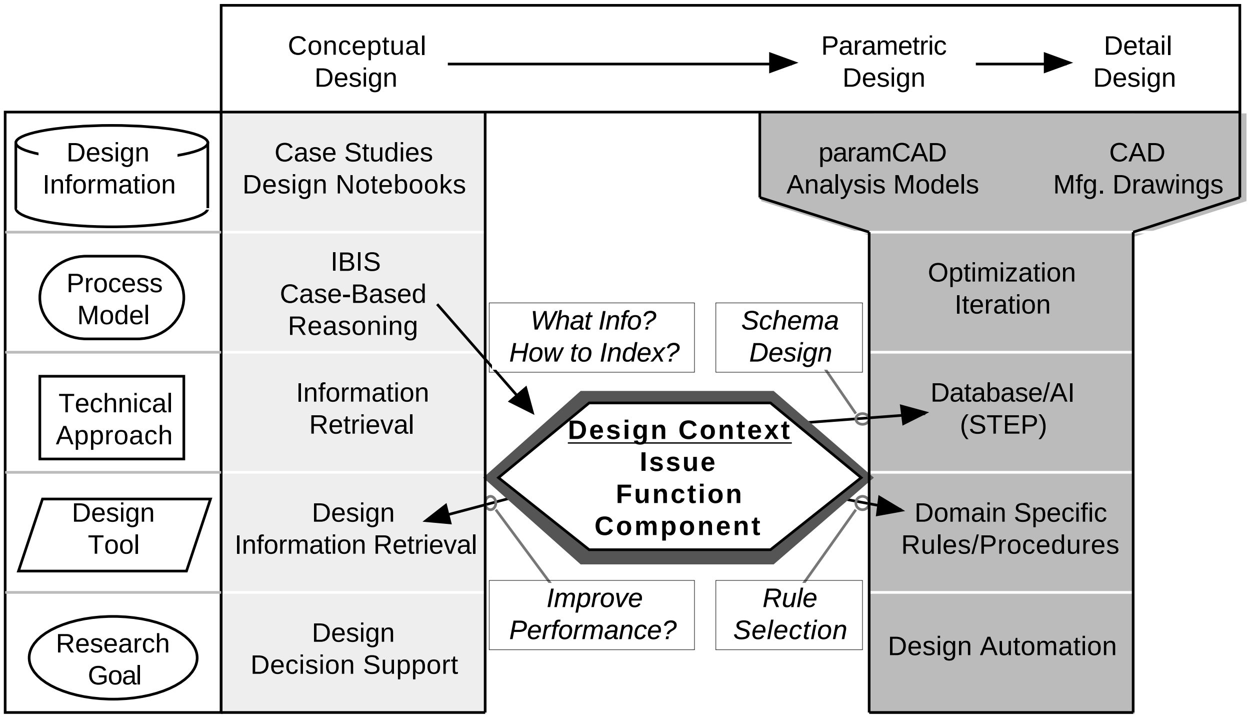

Figure 1 Informal and formal strategies for supporting (respectively) early and late stages of design both depend on formalization of design context. Conceptual design is targeted for information retrieval techniques. Parametric and detail stages might take advantage of formalization through database or knowledge-based methods. Design context is a primary measure of design situation similarity whether it is used for information retrieval, database querying, or knowledge-based rule selection.

The remainder of the paper describes the preliminary steps toward providing an effective and usable method for retrieving informal, unstructured information from the design process. We first describe the nature of information in the design process, specifically focusing on the earliest phase of the design process - conceptual design. Then, a brief background in text-based information retrieval is given. An information retrieval method based on expanding user queries through design thesauri (a generic one for design issues, one specific to design instances) is presented along with experimental results. The implications of this research and avenues for its extension complete the presentation.

2 RELATED RESEARCH

While we have used the term *information* generically, our concern is primarily with *design information*. Thus, a brief discussion of both the types of information generated during the design process and the ways in which this information might be exploited toward improving future design endeavors is appropriate. Because we aim to develop a generic design information retrieval method, the classification of design information in terms of the generality with which it might be reused is significant. Figure 1 illustrates the type of information prevalent during different stages of the design process. Conceptual design information is presented as informal notebooks or case experience, the design process is one of negotiation. We propose to apply information retrieval techniques toward creating a generic *design* information retriveal tool to support informal information gathering for decision making. Parametric and detail design deal with much more concrete information in an iterative optimzation process. Thus database and AI techniques are applicable toward automating the design process. In either case, capturing design context is an essential component. It is hoped that modeling design context, the primary task underlying design information retrieval, will provide the crucial link between informal and formal methods.

2.1 Information and the Design Process

We must begin our discussion of design information retrieval by first settling on a notion of information in the context of the design process. Perhaps the most generic view of the design process comes to us from observations of design practice in the field of architecture. Kunz and Rittel (1970) developed the Issue Based Information System (IBIS) as a process model for design based on negotiation, identifying three main components:

Issues - Communicate the concerns of various stakeholders
Positions - Represent design alternatives
Arguments - Evaluate alternatives with respect to issues

It is important to note that IBIS brings some formalization to the negotiation process because discourse must be labeled. This discourse, however, is not a strict flow from issue to position to argument - often developing a position or an argument spawns a new issue to be resolved. Negotiation in IBIS brings consensus among a team of individual design stakeholders; the set of prototypes which engineers would typically call 'the design' arise as a side effect of this process. Bucciarelli (1988) points out that this notion of 'the design' does not capture the ambiguity that "persist(s) because the design is a socially held entity". Basically, the *real* design is made up of the interrelated concerns of the stakeholders (the design *context*) and how they shape exploration (i.e., identification and evaluation of design alternatives) of the design space.

Because the IBIS model for the design process is derived from descriptive research, it is reasonable to expect its elements to appear in design even if its methodology is not explicitly followed. In fact, many efforts at design process formalization fit into the IBIS framework. Redux (Petrie, 1993) is the most explicit of these, capturing issues as goals and, based on an evolving set of arguments, prompting reconsideration of positions taken. Intelligent Real-Time Design (Bradley and Agogino, 1994) describes a decision and information value theoretic framework which raises issues to help the design team choose among specific alternatives. Wood and Agogino (1997) extend this methodology into a framework that explicitly deals with design abstraction, directing the attention of the design team by analyzing the design space to order design issues according to the team's stated objectives. Ullman (1994, Nagy et al., 1993) extends the IBIS framework to encode the resolution of issues directly in terms of decisions. The process model underlying the Dedal system (Baudin et al., 1993) also includes issue, alternative, evaluation, and decision. A more formal, STEP-based representation for capturing and tracking design positions and arguments within an implicit process model of negotiation and iteration has been proffered as well (Qureshi et al., 1997). These are just a sample of efforts which explicitly deal with the information generated by and supporting the design process, but the breadth of activity leads one to conclude that the information types from IBIS generically represent the design process. Issues, positions, and arguments are thus the focus of development for our design thesauri.

2.2 Reusing Design Information

With IBIS enumerating the types of information present in the design process, we now concentrate on reusing this information. Case-based reasoning (CBR) has been studied by the artificial intelligence (AI) community for the application of experience to current tasks. Operating on information bases smaller than practical for statistical or neural learning (Szykman, 1996; Peplinski et al., 1996, Wood and Agogino, 1997) and with too little formalized knowledge of the domain for single instance or explanation-based learning (e.g., Segre, 1987), CBR seems the perfect fit as a reuse strategy for design. In fact, it has been applied extensively in the past with applications in structural engineering (Maher et al., 1995), building layout (Garcia and Howard, 1992), system design (Navinchandra, 1988), and aircraft design (Domeshek et al., 1994).

The most generic description for CBR is simple: find similar cases, learn what you can from each one, and synthesize a new solution based on what you learned. The design implementations above run the gamut of formality of information from strict, predicate logic based reasoning (Maher et al., 1995), to abstract qualitative reasoning (Navinchandra, 1988), to design objective rationality diagnosis (Garcia and Howard, 1992). Having applied a series of successively less formal CBR techniques in the domain of conceptual architectural design (Pearce et al., 1992; Domeshek and Kolodner, 1992), Domeshek and Kolodner provide the following insight about the storage of design case information (Domeshek and Kolodner, 1993) (*our* underlines):

1. Organize cases into short, pointed presentations that teach specific lessons based on particular experiences.
2. Index such stories in terms of <u>design situations</u> they address.
3. Describe design situations in terms of <u>design issues</u> associated with particular structural or functional parts of an <u>artifact</u>, and remember to consider issues arising from all phases of the artifact's <u>life-cycle</u> from the points of view of all relevant stakeholders.
4. Explicitly note <u>interaction between design issues</u> to broaden the user's focus and draw their attention to related aspects of a design with which they should be concerned.
5. Link stories of specific successes and failures to general guidelines which in turn link back to other related stories in order to allow the user to easily <u>explore a range of responses to the same basic issue</u>.

Of particular note here is the nature of design information that is useful. The emphasis is not on machine understandable codes but rather on human-interpretable information. In addition, considerable retrospection may be needed to discover and contextualize lessons learned from design instances. Beyond retrospection on the importance of issues and the success of choices, design information must also be rooted in a rich description of its context within an artifact description. Petroski (1985) seconds this view by pointing out notorious engineering failures that resulted from a lack of rationale capture in formal documentation. Nor is the artifact itself capable of communicating design issues of secondary importance which might dominate its extrapolation (Petroski, 1994).

The importance of informal information for both capturing and communicating design knowledge leads us to study two (textual views of multimedia) collections of design information in this work: (1) case studies of industry best practices and (2) design notebooks. The former represents a high-level, considerate treatment of interrelationships among life-cycle design issues in a set of retrospective case studies (Agogino and Hsi, 1994). The latter consists of entries that document the evolution of an artifact including all of the alternatives, evaluations, and decisions associated with its design. Studies of these two types of collections are presented in the methods section of the paper. In these sections, we also examine the use of design context in aiding retrieval. We define design context as a mapping into conceptual clusters in hierarchical design thesauri and use this mapping as an intermediate representation for improving retrieval of informal design information. Two such thesauri are studied: one representing life-cycle design issues, the other encoding specific designs in terms of structural and functional decomposition and relating these to generic components and functions.

2.3 Background: Design Information Retrieval

Having described what we mean by design information and context, we now provide a brief introduction to information retrieval (IR). The two basic components of IR are documents and queries; the goal is to match the best documents from a collection (i.e., a corpus) to a given query. To do this, documents and queries are transformed into vectors containing:

Terms - Each word in a document is mapped into a symbol called a term. Each term found in the document corpus is represented as a vector index, the total number of different terms determines the length of the query/document vector.

Term Weights - At each vector index, a real value measuring the degree to which the corresponding term is present in the document.

With both queries and documents represented by vectors, a dot product determines those documents most closely aligned with the query. Some additional tweaks can be applied to this method:

Inverse Document Frequency (*idf*)- The *idf* is a measure of how a word is distributed within the corpus (usually the inverse of the fraction of documents containing a term). *Idf* is used to scale the vector space, shrinking it along dimensions representing commonly occurring words.

Stopwords - Very common words are often removed from the term list entirely.

Stemming - Words with common roots can be mapped to the same term.

Synonyms - Words that are strict synonyms can be mapped to the same term.

Table 1 (Salton, 1988)

Indexing Method	Impact on Performance
Basic single-term automatic indexing	benchmark
Use of thesaurus to group related terms in the given topic area	+10% to +20%
Use of automatically derived term associations obtained from joint term assignments found in sample document collection	-10% to 0%
Use of automatically derived term phrases obtained by using co-occuring terms found in the texts of sample collections	+5% to +10%
Use of one iteration of relevance feedback to add new query terms extracted from previously retrieved relevant documents	+30% to +60%

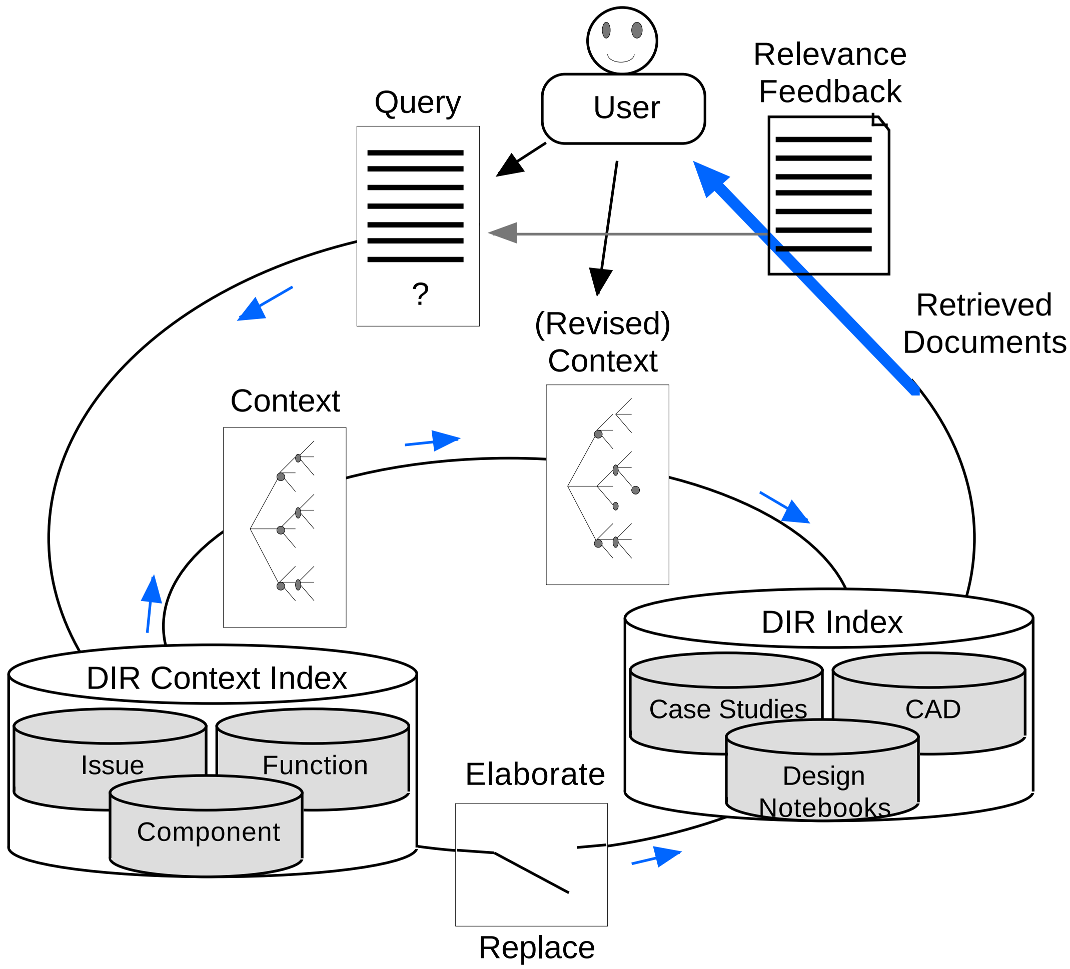

Figure 2 Flow of Information in DIR System. User Query generates context, which can be edited by the user. Context can then replace or elaborate original query when searching design information database.

Word Position - Terms that appear close to each other in both the query and the document may be weighted more heavily than terms that are widely separated.

Noun Phrases - Noun phrases can sometimes be lumped as a single term in vector space.

In Table 1, Salton (1988) describes some empirical results from applying various of these improvements to information retrieval over generic corpi. Most significant among these results is the dominance of 'relevance feedback' (i.e., applying a set of relevant documents as the IR 'query') as the best means for enhancing system performance. This is born out by more current research in the IR community: the Text Retrieval Conference (TREC) (TREC) offers a standard set of information retrieval contexts and set of documents from previous years whose context relevance has been assigned (none was available for the first year). Using relevance feedback to elaborate standard queries provides a tremendous increase in system performance over ad-hoc querying. By limiting our focus to design information retrieval and using specially tailored thesauri, we hope to better the 10-20% performance improvement experienced by Salton. We will additionally relax the mode of thesaurus implementation from that of strict synonymy to allow query elaboration with multiple word 'terms', acronyms, and descriptive phrases.

Figure 2 shows the overall flow of information in the design infomation retrieval (DIR) system. The user formulates a query which is then matched against the DIR context index (i.e., the design thesauri) to determine its *design* content. This machine-generated context can then be edited by the user for accuracy. This design context information can elaborate or replace the original query when presented to the search engine operating on the actual design information. In addition, retrieved design documents can be used to generate relevance feedback additions to the original query and contextualization. Of note are the two significant points of user feedback in the DIR system:

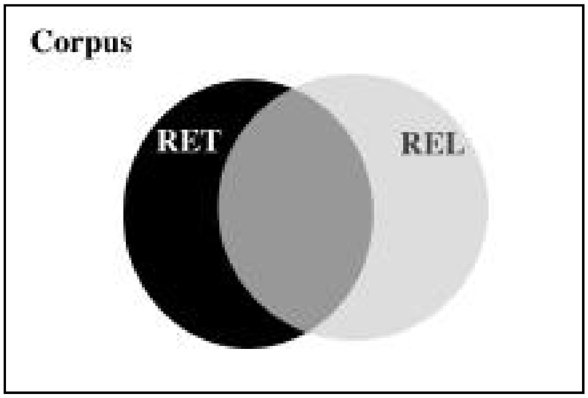

Figure 3. Venn Diagram of Retrieval

correction of the machine-determined design context and relevance feedback. System parameters of experimental interest are the determination of replacement/elaboration strategies and the breadth of design context used. In order to understand the impact of these variables, metrics for measuring IR performance must be discussed.

Two closely coupled measures are used to assess the performance of an IR system: precision and recall. Figure 3 is a Venn diagram in which the document corpus is the universal set. In response to a user query, two subsets are generated: the retrieval set (RET) representing the documents from the corpus which the system has identified as being relevant to the user query, and the relevant set (REL) which is the set that contains all of the documents actually relevant to the query. The system performs perfectly when these subsets are identical. The two performance metrics for a query can be calculated once RET and REL have been identified:

Precision = #of docs in (RET ∩ REL) / #of docs in (RET)

Recall = #of docs in (RET ∩ REL) / #of docs in (REL)

The basic research method for determining IR performance is: develop a representative query set for the information base, determine REL for each query (usually through consensus among human experts), and get the system to generate RET for the various IR strategies under consideration. Precision and recall are derived for each strategy to determine which are most effective. As was mentioned, these two measures have empirically been determined to be coupled according to the following relationship:

Precision*Recall = **K**
where **K** is a system constant

Figure 4 illustrates this relationship and illustrates the goal of our DIR experiments: to determine the search strategies which maximize **K** while allowing precision to be traded against recall according to information gathering needs. We will

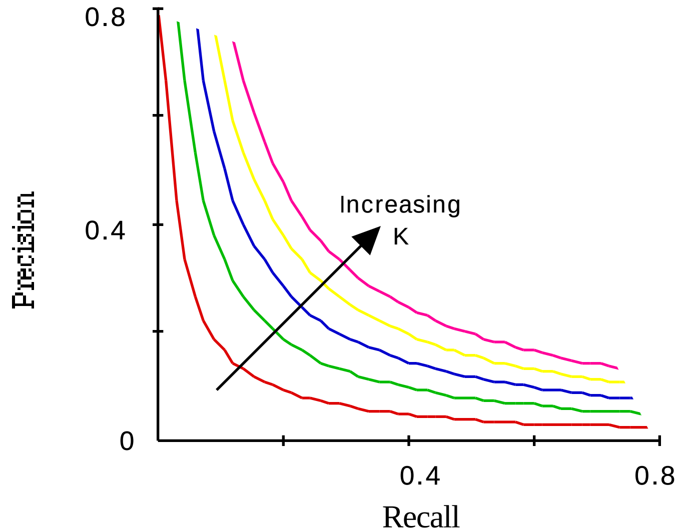

Figure 4. Empirical Relation Between Precision and Recall. An increase in the combined performance measure, **K**, provides better system performance, reducing neither Recall nor Precision.

address the information reuse concerns discussed above by studying two distinct information bases - a set of life-cycle design case studies which test design issue indexing and a set of design notebooks to test function / component indexing. The common thread tying these efforts is the development and use of a design thesaurus for representing design context. We will now separately discuss the issue thesaurus and its performance followed by the component/function thesaurus.

3 INDEXING BY DESIGN ISSUE

Wood and Agogino (1996) describe the development of a hierarchical set of issues generic to life-cycle design. Containing over 1,000 points derived from analysis of many sources (e.g., Boothroyd et al., 1991; Deiter, 1991) and related to each other by human experts, the hierarchy represents distinct design concepts with a set of descriptive, synonymous phrases and a brief definition containing typical related issues. Figure 5 demonstrates thesaurus coverage in terms of breadth and typical depth. Figure 6 shows a typical thesaurus item.

3.1 Issue-Based Retrieval Experiments

A series of experiments were conducted to assess two primary system design considerations: how to assign design issue context to a query by mapping it into items in the issue thesaurus and how to use this context to improve design information retrieval as measured by precision and recall. The experimental methodology is as follows:

Queries: Relevance feedback querying was chosen as the

- Life Cycle >
  - Design >
    - Concept Generation >
      - Brainstorming >
      - Issue Identification >
      - Decomposition
    - Service Conditions >
      - Load >
      - Environment >
    - Dynamics >
    - Statics >
    - System Reliability >
    - Failure Mode >
      - Environmental >
      - Wear >
      - Fatigue >
      - Plastic Deformation >
      - Elastic Deflection
      - Load Instability >
      - Combination >
      - Impact >
  - Marketing >
  - Distribution >
    - Sale >
      - Packaging >
      - Pricing
    - Delivery >
  - Manufacturing >
    - Forming Processes >
    - Assembly Processes >
      - Fastening >
      - Fixturing >
      - Time - Motion
      - Part Count >
      - Part Insertion >
      - Part Handling >
      - Assembly Line >
      - Assembly Method >
    - Process Control >
    - Finishing Processes >
      - Cleaning >
      - Coating >
  - Service >
    - Maintenance >
    - Repair >
      - Diagnosis >
      - Repair Action >
      - Diagnosability
    - Upgrade >
  - Retirement >
    - Waste Stream
    - Recycling >
      - Disposal

Figure 5 Partial Design Issue Hierarchy. Shown in outline form for clarity, the issue hierarchy expands further at each leaf demarked with a “>”. Over 500 nodes are used to contextualize design issues.

current best method for querying and to eliminate possible bias in the query set. Full documents were extracted at random from the Base Corpus and offered to the panelists as the basis for retrieval relevance evaluation. This test is one of ‘virtual hypertext’ where the goal is to determine which links are most intersting given the current document context. It also removes testing bias that might be introduced by specific queries.

Base Corpus: The corpus of materials from which query documents was drawn is a collection of multimedia case studies described in (Wood and Agogino, 1996). The text from each case study page was extracted, indexed in an IR system (WAIS (Kahle et al., 1993)) and referenced back to the originating multimedia page. The total size of the collection is about 500 pages.

External Corpus: A mechanical engineering handbook (Kutz, 1986) of approximately 2000 pages was scanned and optical character recognition performed

Term: Assembly
Synonyms: Put together
Description: Construct subsystem by putting parts or components together.
Parent: Manufacturing
Children: Fastening , Fixturing , Time - Motion, Part Count, Part Insertion, Part Handling, Assembly Line, Assembly Method, Process Control

Figure 6 Design Issue Thesaurus Item. Slots include parent(s), term, synonyms, descriptive phrase, child(ren).

on each scanned page (uncorrected, 'dirty' OCR). The resulting text was indexed through an IR (WAIS) system and associated back to the original page image.

A panel of five design experts with varying familiarity with the collection of case studies was chosen. Provided a set of 40 base documents, each expert was asked to evaluate retrieval episodes for the ten closest to his design expertise. Panelist performed three tasks for each base document: mapping of the document into the design issue thesaurus, judging relevance of documents retrieved from the Base Corpus, and judging relevance of documents retrieved from the External Corpus.The subset of corpus documents over which relevance was judged for each query was pooled from high-recall results sets from all experimental system configurations. Results from the three tasks follow; interpreting each legend key determines the content of the query:

T: The context term is added to the query
Xt: Compound context terms are matched exactly
S: Term synonyms are added to the query
Ph: The descriptive phrase
P: Term parents are added to the query
C: Term children are added to the query
stem: The database was searched using stemming.

For information retrieval results, the following prefixes or substitutions appear:

Q: The original query is used (i.e., query elaboration). Absence of Q indicates query replacement.
IR: The original query is used alone.
Auto: Automatic context assignment is used (the trailing number indicates the WAIS threshold above which context terms are added)
Corr: Context assignment is corrected by user from among automatically assigned set.

3.1.1 Experiment 1 Design Issue Context Assignment Results

Figure 7 shows a partial result set derived from the contextualization of case study documents into the design issue thesaurus. The relevance judgments of the experts were used to benchmark the performance of a variety of information retrieval strategies. In this case, the document was passed as a query over a series of collections of documents created by representing each thesaurus concept with a varying amount of information. Some comparisons are in order:

Stemming: Note the difference between the performance of 'TS' (the term+synonym database) and TS-stem (the same database searched using stemming). For a given precision of result, searching using word stems produces a greater recall. However, the maximum precision that can be derived from a search is found by searching the unstemmed database.

Context Specificity: The lines 'TS - stem', 'TSPhPC - stem', and 'TSPhPCPh - stem' (decreasing in context specificity respectively) show a classic precision vs. recall tradeoff. As a design issue is

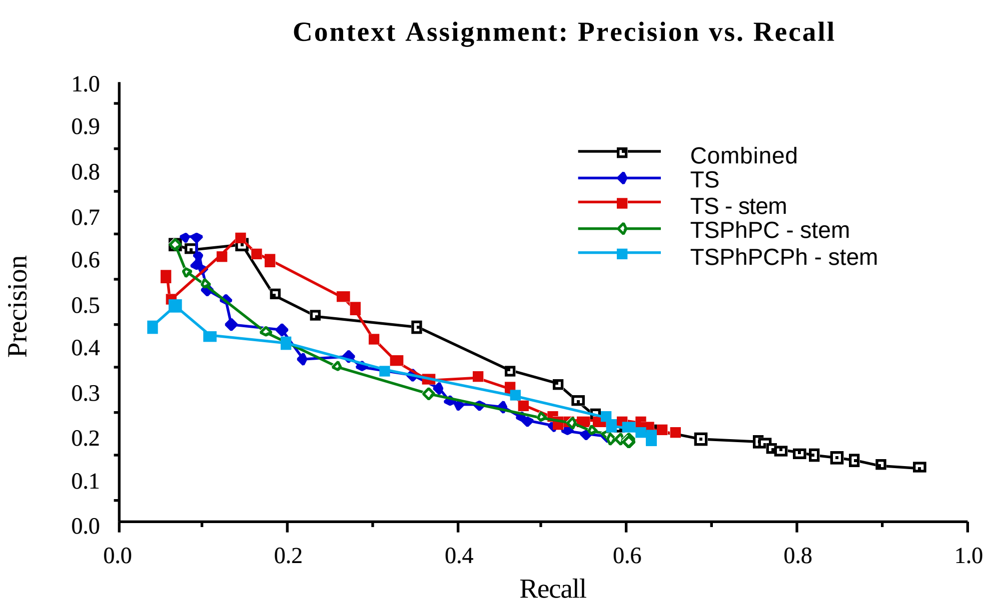

Figure 7: Contextualization of Case Study Documents

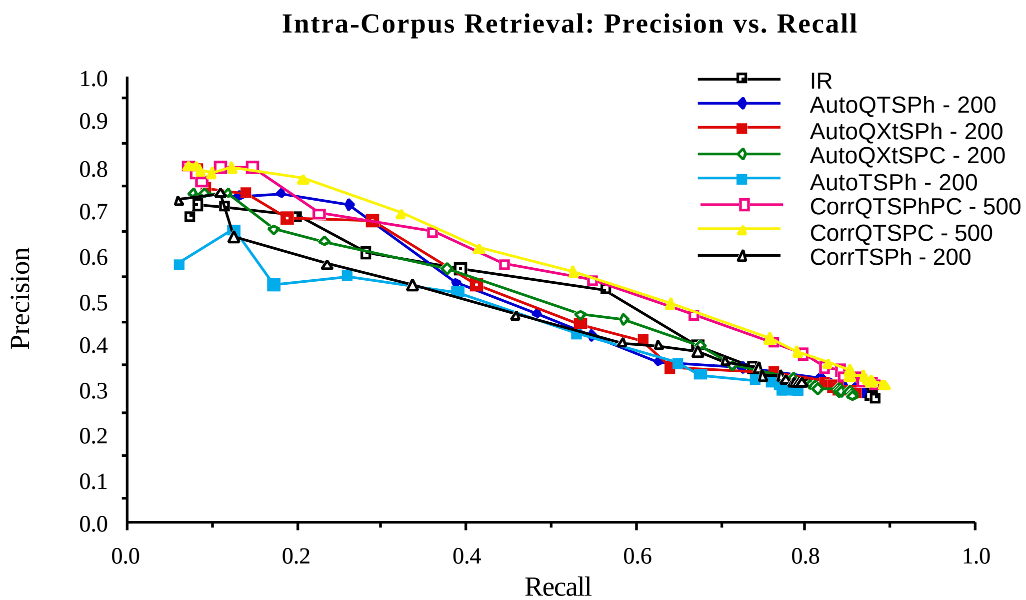

Figure 8: Query elaboration/replacement within the base corpus

described in broader terms, recall improves (a flatter precision vs. recall curve) at the cost of precision.

These two points summarize one of the design points within the system: without specific knowledge of a user's preferences between precision and recall we cannot determine a best strategy for stemming or for context specificity. The graph labeled 'Combined' represents a simple attempt (combining all tested techniques into a composite one) to provide an example of how the 'sum is greater than the parts'.

3.1.2 Experiment 2 Issue-Based Retrieval Results: Intra-Corpus

Figure 8 compares the performance of various query elaboration/replacement techniques for information retrieval within the base corpus. Of note here is the baseline technique of relevance feedback, labeled 'IR'. Plots prefixed with 'A' represent automatic contextualization and inclusion of thesaurus items according to the codes from the previous section. Plots prefixed with 'C' represent a correction by the user of the thesaurus context. Plot strings containing 'Q' include the original query, otherwise it is excluded. Two points of comparison are to be derived from this plot:

Query Elaboration vs. Query Replacement: Plots 'AutoQTSPH - 200' and 'AutoTSPH - 200' demonstrate the importance of the original query within the same corpus of information. The former run, adding terms from the thesaurus to the original query document, attains higher precision throughout the range of recall compared with replacing the original query with the same terms

Automatic vs. Corrected Contextualization: The general trend shown in the figure is a precision improvement of 10-30% in favor of a single step of context assignment feedback from the user. This is significant search improvement at the cost a couple of simple mouse clicks to remove spurious thesaurus terms.

In summation the design thesaurus improves query performance, but not markedly when compared to the baseline relevance feedback. This is perhaps an artifact of the common jargon used across the corpus which is not replicated by a thesaurus designed for generality. However, this thesaurus is relatively successful in capturing design context as demonstrated by the small difference between using a thesaurus-derived set of concepts for querying in place of the base document relevance feedback. We will now consider how these results extend to retrieving design information from outside of the original document collection.

3.1.3 Experiment 3 Issue-Based Retrieval Results: Extra-Corpus

The set of experiments for extra-corpus information retrieval begins to demonstrate the true value of using a design issue thesaurus. Figure 9 again uses IR to represent baseline relevance feedback for querying the information base (in this case, an asterisk (*) added to the legend string denotes the use of stemming). We now revisit the performance comparisons of the previous section:

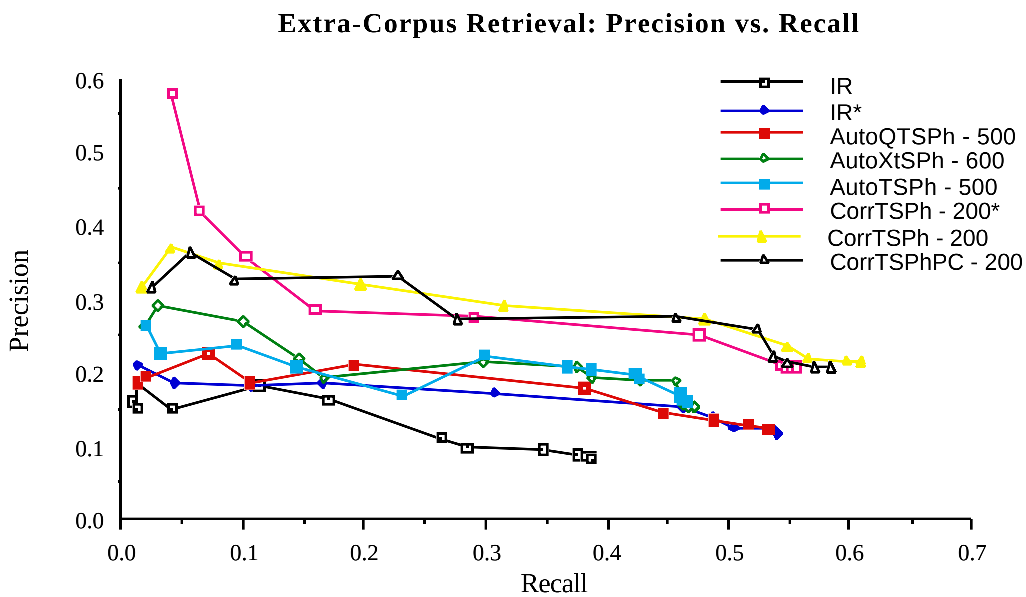

Figure 9: Design issue thesaurus on extra corpus information

Query Elaboration vs. Query Replacement: Plots 'AutoQTSPh - 500' and 'AutoTSPh - 500' demonstrate that in a heterogeneous information base, it is more effective to replace the original query with its design context than to reinforce it with terms from the thesaurus. Either method performs much better than simple relevance feedback (again labeled IR).

Automatic vs. Corrected Contextualization: Instead of 10-30% improvement derived from correcting design context we now see on the order of 100% precision improvement for the same recall levels.

We would be remiss if we did not note the absolute decrease in query performance compared to intra-corpus retrieval. In the above experiments, querying external to an information base is only half as effective as querying internal to it. It is suspected that this is due, in large measure, to the relative conceptual proximity of documents within the base corpus. However, the performance improvement gained by interposing a thesaurus between information bases is clearly demonstrated.

Overall, the results of the implementation of a design issue-based thesaurus are quite promising. We are able to routinely achieve recall levels of approximately 100% or precision levels of up 80% (and many compromises in between). Such performance is almost to be expected when searching design information which is 'designed' to be most generic - case studies and design handbooks. We now turn our attention to the potentially more difficult problem of improving query performance within actual design documentation.

4 INDEXING BY COMPONENT/FUNCTION

The previous section presents an approach to retrieving design information based on the use of generic design issues to capture design context. In this section, we examine another source of design context: a component/functional relation model of the design itself. Design notebooks created during the design process itself are very different in scope from the case studies and design handbooks discussed above. Case studies are a distilled version of information from the design process, formatted into a cohesive story, while design notebooks archive design process information, warts and all. Again, the approach is to determine what is generic about retrieving design information form these notebooks. In addition, we must understand how we might customize of the retrieval process for eachn notebook and the tradeoff between the effort required for customization and its impact on retreival performance. To this end, we create component and function thesauri (i.e., models of the artifact being designed in these notebooks) derived from two sources: the actual notebook (high effort) and final design documents (lower effort). These thesauri are tested in a desgin information retrieval framework based on the Dedal system (Baudin et al., 1993), extended to take advantage of information retrieval techniques.

4.1 Introduction to Dedal

From analysis of design protocol studies, (Baudin et al., 1991) devised a two part framework for engineering queries. These parts are the <descriptor> and the <subject> (Figure 10). The <descriptor> is a generic engineering concept that crops up repeatedly in design discourse. Engineers want to consider *alternatives*, for example, and examine *assumptions*. The

Descriptor	+	Subject
<Alternative>		*Part of an object model*
<Assumption>		*Ex:*
<Comparison>		
<Construction>		
<Location>		
<Operation>		
<Performance>		
<Rationale>		
<Relation>		
<Requirement>		

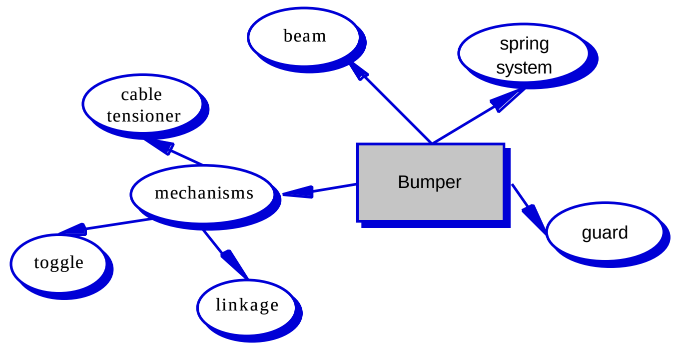

Figure 10: The Dedal Framework

<subject> is a specific part of the device model, such as a motor or a linkage. Together, these form a Dedal query.

In previous incarnations of Dedal, final report documentation was hand indexed for <descriptors> and <subjects>. This resulted in precision of up to 70% and recall up to 90%, somewhat lower than values for design issue indexing. However, manual indexing requires a great deal of overhead. An indexer has to be both familiar with the subject matter (presumably an engineer on a project) and have the time to mark all relevant text.

DedalAI strives to capitalize on the power of the Dedal's design information retrieval model while mitigating the resource demands of hand indexing design text. Application of a generic thesaurus for the descriptor query component was explored by Yang and Cutkosky (1997). This thesaurus was constructed by hand, using terms drawn from both general purpose thesaurus and the text of design notebooks themselves. While this is a time consuming process, the thesaurus does not have to be continually generated because of the generic, stable nature of descriptors. An example synonym for the descriptor <Alternative> would be "possibilities". In this case, the Dedal query <Alternative> of <actuator> would return any chunk of text containing both the word "possibilities" and the manually indexed subject <actuator>. Tests on three different electronic notebooks with three different descriptors improved retrieval precision between 30 and 50% over non-thesaurus searches. Test for breadth of application of the descriptor thesaurus proved its utility on other types of documents, such as patents, with similar content.

With a viable descriptor thesaurus in hand, we now turn to the development of a subject thesaurus to complete the DedalAI framework. Using a similar approach to that taken descriptors two sources for thesaurus terms present themselves: final project reports (including CAD drawings and diagrams) and informal project design notes. The formality of the final reports eases the task of creating a thesaurus, but because these reports are generated specifically due to the academic nature of the design projects they may not generalize to non-academic design (although certainly the CAD portions of final reports are generic). On the other hand, the design notebooks represent the generic communication that takes place in the process of team-design. The question is: What is the better source of information for generating a subject thesaurus, informal in-process documents or formal final design documents? The experimental methodology for exploring this question is outlined below:

Queries: In Dedal (and DedalAI), plain English questions are translated[1] into <descriptor> <subject> query format. For example, the question "What actuator alternatives were considered?" becomes the query *<Alternative> of <actuator>*. Generic design queries were created based on experience of actual design information demands (Baudin et al., 1991).

Corpus: Three sets of project documentation, including the electronic design notebooks from all designers and the final report, from a graduate level course in electromechanical design. Students work in teams of three or more over a nine-month period, starting from conceptual design to working prototype. The projects examined here include the design of a car bumper, a fountain, and a personal digital assistant. These projects are representative of the ten or so projects in the class, and also cover different types of design: redesign of an existing product and new conceptual design. Typical corpus size is ~5000 documents (~2MB of ASCII text).

[1] Untranslated queries were studied and provided quite poor results, automatic contextualization like that done for design issues is not covered in this paper.

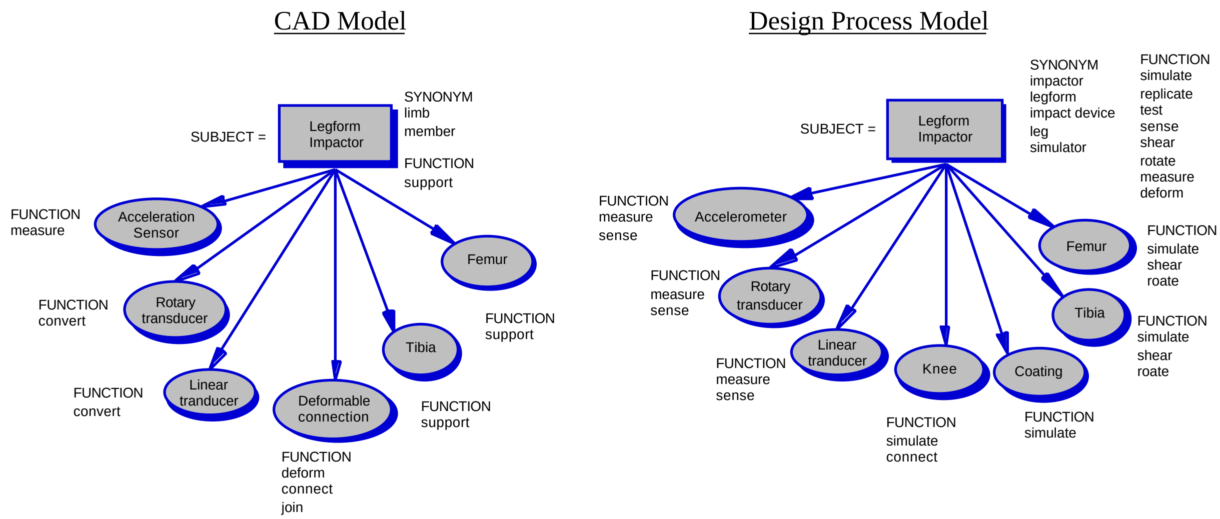

Figure 11: Parent-child clusters based on CAD drawings and design process documentation

The electronic design notebooks were created in PENS (Personal Electronic Notebook with Sharing) (Hong et al., 1994), a tool for generating collaborative, Web-based notebooks from text and images. Students were strongly encouraged to document as much of their work as possible in PENS, from to-do lists to formal reports. As a result, the PENS notebooks document much of each team's design process. Notebook entries vary widely in nature, from fragmentary to well-organized. Comprehensive final reports for the class were generated by teams at the end of the year. Reports contained detailed information on the final design, such as CAD drawings and diagrams, as well as content drawn from PENS notes.

4.2 Model Building and Thesaurus Generation

Two artifact models were created for each project, one drawn from information found in final reports (we will label these CAD models), another from design model fragments found throughout the design notebooks (labeled DP models). The level of formality of these models is sufficient only for determining system decomposition and for assigning names to subsystems. Synonyms for both form and function were assigned to each node. Examples of CAD and design process models for the car bumper project are shown in Figure 11.

Synonyms and function for parts of the CAD model were found by using a Web-based dictionary and thesaurus engine (Hypertext Webster Gateway at UCSD). This dictionary includes searches on Webster's Dictionary and WordNet, a semantic net thesaurus, and is presumably an objective, repeatable way of finding such information. While searching for these synonyms and functions, however, it quickly became obvious that many of these terms, like "Pedestrian Impact Guard", are too specific to be found in a general purpose dictionary. In these cases, the most general form of that part name, such as "guard," was used.

Creation of models for the DP thesaurus was less straightforward than for the CAD models because of the implicit and incomplete nature of model description in something informal like PENS notes. While some diagrams and drawings are provided in PENS notebooks, they aren't contextualized as thoroughly as might be found in a final report. Ideally, neat models could be extracted from a notebook from time to time to show "snapshots" of a changing design, but it is difficult to create a single model for an evolving design. For this reason, fragments from various stages of design are linked together (e.g., a bumper and impact guard are composed).

Form synonyms and functions for the DP model were generated by examining the design notebooks carefully for references to the part. Two distinct ways of referring to parts were found: 1) domain specific synonyms and 2) generic design words. Domain specific synonyms for the supporting structure of the fountain are "skeleton", "box", or "cradle"; their association requires some understanding of the project and its jargon. Generic design words were used like pronouns to describe a design: the fountain project's nozzle is variously called a "system," a "device," a "prototype," and a "design."

4.3 DedalAI Retrieval Results

Retrieval runs were performed for a set of questions for each design notebook. In each case the descriptor thesaurus was used to add synonyms to the query corresponding to its descriptor element. Various strategies for applying the subject thesaurus are given below, organized by their label for Figure 12:

SubjOnly: No thesaurus, subject alone
SubjArt: Add form synonyms for the subject (artifact)
SubjFunc: Add function terms for the subject
SubjArtFunc: Add both form and function synonyms
PC: Add parent/child terms

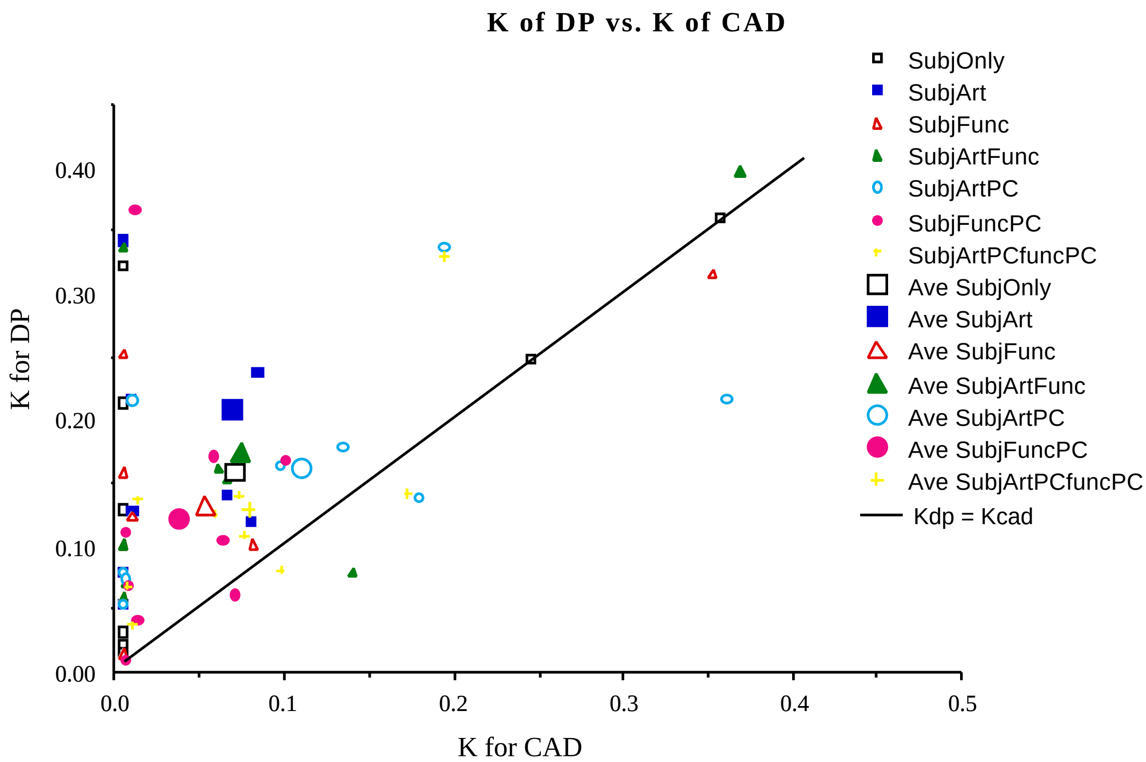

Figure 12: Comparison of DP and CAD thesaurus by K value

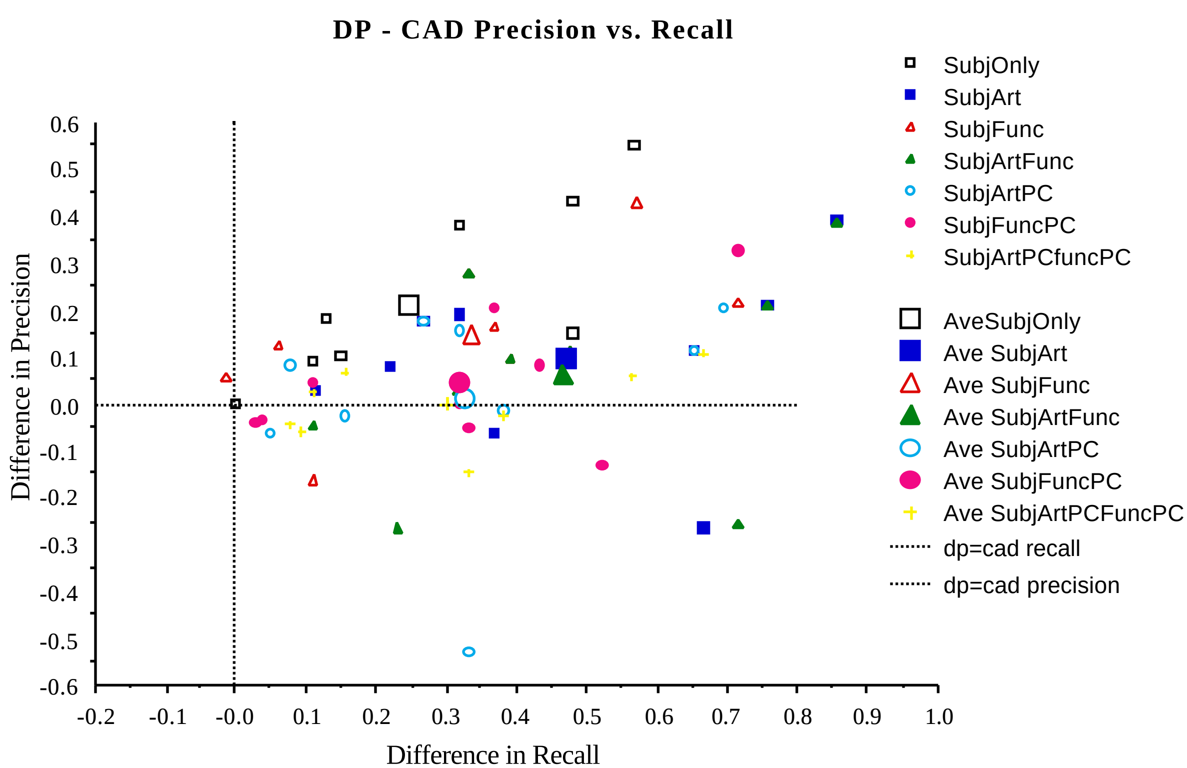

Figure 13: Comparison of Precision and Recall Differences for DP and CAD thesauri

As stated above, the product of precision and recall (K) roughly indicates the overall performance of a particular information retrieval approach. Figure 12 compares K values for the design process thesaurus against those for the CAD thesaurus for each retrieval episode. Because the majority of points fall above the Kdp = Kcad line, it appears that the DP thesaurus is more effective at finding correct answers than the CAD thesaurus. Of particular note is the set of points aligned near the vertical axis. This indicates that the design process thesaurus is effective in cases where the CAD thesaurus is almost completely useless, a strong result which might indicate a change in language from the design process to the final design documentation.

Figure 13 details the results of Figure 12, showing the difference in precision and recall performance between the two thesauri. The design process thesaurus almost always out performs the CAD thesaurus with respect to recall and generally improves on its precision as well. Again, if one wants access to in-process design documentation it appears to pay to align queries with the less-formal language found there. The largest average gap between the two thesauri comes when only SubjArt or SubjFunc terms are added, showing a 50% difference in recall with a 10% improvement in precision.

4.4 Discussion of Function/Component Indexing

The generation of models from CAD drawings and other diagrams from a final report are straightforward. Part names and relationships are relatively unambiguous. In PENS design notebooks, informal, partial device models are generated constantly throughout the design process. These models are usually fragmentary, with the team concentrating on only a portion of the design at a time. Figure 14 shows how the view of a design that emerges from final documentation differ substantially from that seen in day-to-day documentation. The language used to describe parts of a design in these notebooks can be very different than the language used in a final report or CAD drawing, potentially chainging with each new design iteration. Immersed in the design task, the language of discourse can also be very general (i.e., calling the fountain assembly the "prototype"). The results of the preceding experiments bear these observations out: To access in-process documentation, there is a clear advantage to accumulating a thesaurus rather than deriving one from the final design.

5 CONCLUSIONS

We have presented a framework for a design information retrieval (DIR) system that is a step towards increasing the reuse of informal knowledge in design. The system utilizes several thesauri, covering design issues, artifact components, and component functions to capture and enhance context for search and retrieval. The results of testing these thesauri on design case studies and design notebooks are summarized below:

- There is no one correct approach to determining the context of a case study for the design issue thesaurus. The choice of strategy depends on a particular user's preference between precision and recall.

- Within the corpus of case studies, the use of a design issue thesaurus can improve query performance compared to relevance feedback systems, though not significantly. The major finding here is that by equaling the performance of relevance feedback, design context comprised of a set of design issues selected from the hierarchical thesaurus adequately represents these documents.

- When the design issue thesaurus is used on outside collections, it provides very promising results. Tests

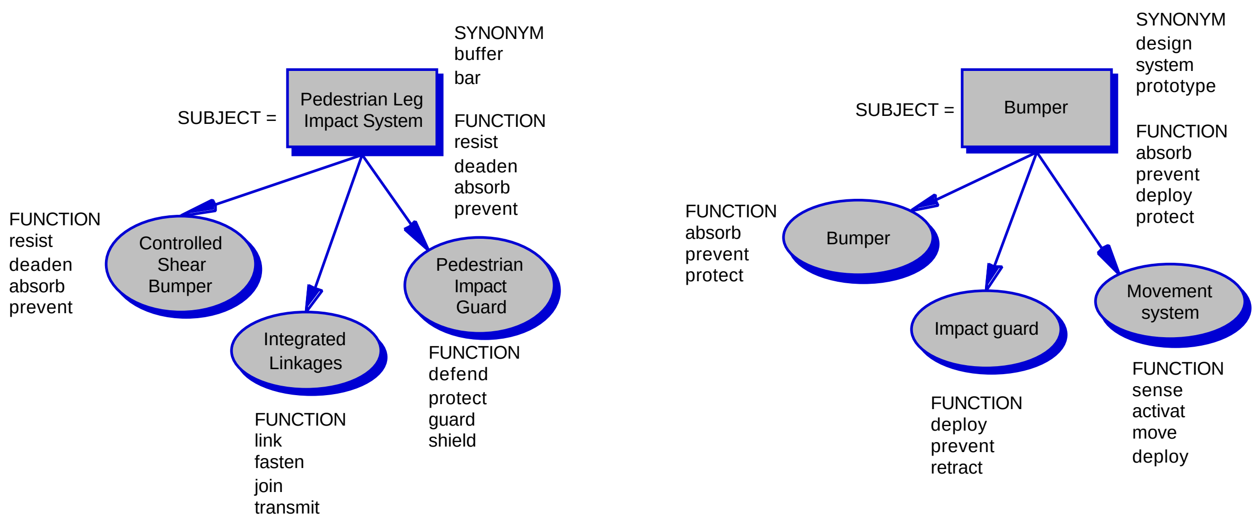

Figure 14: Difference in part naming in CAD and DP models

consistently show recall levels of 100% and precision of nearly 80%. Thus the generality of the issue thesaurus is demonstrated.

- The use of a thesaurus of function/component terms derived from design notebooks performs better than a thesaurus derived from formal final reports. The trade-off lies in the effort required to create the models. Models based on final reports are straightforward to derive, while models from design notes require some knowledge of the domain.

These results suggest that the DIR approach of exploiting design context as a way to search and retrieve design information will eventually lead to a viable system for helping designers gain access to external, informal design knowledge, the first step towards its effective reuse.

Future plans for this work include using automatic methods to generate subject thesauri. One method currently being explored is the application of singular value decomposition (SVD) to terms in the PENS notebooks in order to extract major themes in the text (Yang, et al, 1998). These themes may then be used as the basis of a thesaurus. Preliminary results suggest that machine methods for generating thesauri are promising, but not quite as effective for retrieving information as relying on hand built thesauri.

In addition, we hope that design context as described above will become a useful tool not only for design information retrieval but also for mediating design communication. Industry is increasingly embracing groupware as a means for fostering teamwork in virtual enterprises. It is not yet clear whether the current generation of communcation software can build the type of community among team members that has proven successful in the form of physical collocation. Design information retrieval based on the notions of design context modeling discussed above, applied to dynamic information streams, may provide the next generation of groupware for design applications.

6 ACKNOWLEDGMENTS

Portions of this work was done under a DARPA DSO-sponsored project, administered with the assistance of ONR, under the RaDEO program and is partially funded by Navy contract N00014096-1-0679. Their support is gratefully acknowledged. In addition we would like to acknowledge the intellectual support provided by Sherry Hsi, the Synthesis Coalition, and the Berkeley case studies group, and the BEST Lab.

REFERENCES

Agogino, A.M., Hsi, S., 1994, "The Impact and Instructional Benefit of Using Multimedia Case Studies to Teach Engineering Design", *Journal of Educational Hypermedia and Multimedia*, Vol. 3, No. 3/4, pp. 351-376.

Baudin, C., Underwood, J. G., Baya, V., Mabogunje, A., 1991, "Dedal: Using Domain Concepts to Index Engineering Design Information", *Workshop Notes from the Ninth National Conference on Artificial Intelligence*, Intelligent Multimedia Interfaces, Anaheim, California.

Baudin, C., Underwood, J. G., Baya, V., , 1993, "Using Device Models to Facilitate the Retrieval of Multimedia Design Information", *Proceedings of International Joint Conference on Artificial Intelligence* .

Beiter, K., Krizan, S., Ishii, K., Hornberger, L., 1993, "HyperQ/Plastics: an Intelligent Design Aid for Plastic Material Selection.", *Advances in Engineering Software*, vol.16, (no.1):53-60.

BDI (Boothroyd Dewhurst Inc.), http://www.dfma.com.

Boothroyd, G., Dewhurst, P., and Knight, W.A., 1991, "Research Program on the Selection of Materials and Processes for Component Parts", *International Journal of Advanced Manufacturing Technology*, Vol.6,No.2, pp. 98-111.

Bradley, S., and Agogino, A., 1994, "An Intelligent Real Time Design Methodology for Catalog Selection", *ASME Transactions, Journal of Design*, Dec., 1994, Vol. 116, pp. 980-988.

Bucciarelli, L., 1988, "An Ethnographic Perspective on Engineering Design", *Design Studies*, July, 1988, pp. 159-168.

Conklin, J., and Begeman, M., 1988, "gIBIS: A Hypertext Tool for Exploratory Policy Discussion", *Proceedings of the COnference on Cumputer-supported Cooperative Work*, ACM, New York, pp. 140-152.

Dieter, G.D., 1991, *Engineering Design: A Materials and Processing Approach*, McGraw-Hill.

Domeshek, E. and Kolodner, J., 1993, "Using the Points of Large Cases", *Artificial Intelligence for Engineering Design, Analysis, and Manufacture*, Vol. 7, No. 2, pp. 87-96.

Domeshek, E.A. and Kolodner, J.L. ,1992, "A Case-Based Design Aid for Architecture", in Gero, J., Sudweeks, F. (eds)*Artificial Intelligence in Design '92*, Kluwer Academic Publishers, pp. 497-516.

Domeshek, E.A., Kolodner, J.L., and Zimring, C.M., 1994, "The Design of a Tool Kit for Case-Based Design Aids", in Gero, J., Sudweeks, F. (eds)*Artificial Intelligence in Design '94*I, Kluwer Academic Publishers, pp. 109-26.

Garcia, A.C.B., and Howard, H.C., 1992, "Acquiring Design Knowledge Through Design Decision Justification", *Artificial Intelligence for Engineering Design, Analysis, and Manufacture*, Vol. 6, No. 1, pp. 59-71.

Hong, J., Toye, G., Leifer, L., , 1994, "Using the WWW for a Team-Based Engineering Design Class", *Electronic*

Proceedings of the Second WWW Conference, Chicago, Illinois, October, 1994.

Hrinyak, M.J, Bras, B., and Hoffman, W.F., 1996, “Enhancing Design for Disassembly: a Benchmark of DFD Software Tools”, *Proceedings of The, 1996 ASME Design Engineering Technical Conferences and Computers in Engineering Conference*, 96-DETC/DFM-1271

Hypertext Webster Gateway at UCSD, (http://work.ucsd.edu:5141/cgi-bin/http_webster)

The Institution of Electrical Engineers, *INSPEC Thesaurus, 1995*, Great Britain, Unwin Brothers Limited.

Kahle, B., Morris, H., Goldman, J., Erickson, T., and Curran, J., 1993, “Interfaces for Distributed Systems of Informatoin Servers”, *Journal of the American Society for Informatoin Science*, Vol. 44, No. 8, pp. 453-467.

Kutz, M. (ed), 1986, *Mechanical Engineer’s Handbook*, John WIley and Sons, New York.

Kunz, W. and H. Rittel, “Issues as Elements of Information Systems,” Studiengruppe Fuer Systemforschung, Heidelberg, Working Paper #131, 1970. (Also described in D. Grant, “How to Use the IBIS as a Procedure for Deliberation and Argument in Environmental Design Planning,” *Design Methods and Theories*, vol. 11, no. 4, pp. 185-220, 1977.)

Lenat, D.B., 1995, “CYC: A Large-Scale Investment in Knowledge Infrastructure”, *Communications of the ACM*, Vol. 38, No. 11, pp. 32-38.

Liang, J., Shah, J., Urban, S., Harter, E., and Bluhm, T., 1997, “STEP Pilot Project: Consolidated Schema for Airframe Engineering Analysis”, *Proceedings of DETC97:, 1997 ASME Design for Manufacturing Conference*, DETC97/DFM-4375

Maher, M.L., Balachandran, M.B., Zhang, D.M., 1995, *Case-Based Reasoning in Design*, Lawrence Erlbaum Associates, Mahway, N.J.

Nagy, R., Ullman, D., and Dietterich, T., 1992, “A Data Representation for Collaborative Mechanical Design”, *Research in Engineering Design*, Vol. 3, No. 4, pp. 233-242, Springer-Verlag, New York.

Navinchandra, D., 1988, “Behavioral Synthesis in CADET, a Case-Based Design Tool”, *Proceedings of the DARPA Workshop on Case-based Reasoning*, Kolodner, J., Ed., Morgan-Kaufman, pp. 286-301.

Pearce, M., Goel, A.K., Kolodner, J.L., Zimring, C., Sentosa, L., and Billington, R., 1992, “Case-Based Design Support, A Case Study in Architectural Design”, *IEEE Expert*, October, 1992, pp. 14-20.

Petrie, C., 1993, ``The Redux' Server,'' *Proceedings of the International Conference on Intelligent and Cooperative Information Systems (ICICIS)*, Rotterdam, May, 1993.

Petroski, H., 1985, *To Engineer is Human: The Role of Failure in Succesful Design* , St. Martin's Press, New York.

Petroski, H., 1994, *Design Paradigms: Case Histories of Error and Judgment in Engineering*, Cambridge University Press, New York.

Peplinski, J., Allen, J., and Mistree, F., 1996, “Integrating product design with manufacturing process design using the robust concept exploration method”, *Proceeding of the, 1996 ASME Conference on Design Theory and Methods*, ASME 96-DETC/DTM-1502.

Qureshi, S.M., Shah, J.J., Urban, S., Harter, E., Bluhm, T., 1997, “Integration Model to Support Archival of Design History in Databases”, *Proceedings of DETC ‘97:1997 ASME Design Engineering Technical Conferences*, DETC97/DTM-3876

Salton, G., 1988, *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*, Addison-Wesley, Reading, MA.

Segre, A., 1987, “On the Operationality / Generality Tradeoff in Explanation-Based Learning”, *Proceedings of the Tenth International Joint onference on Artificial Intelligence: IJCAI ‘87*, IJCAI, Los Altos, CA.

Szykman, S., 1996, “Improving the design process by predicting downstream values of design attributes”, *Proceeding of the, 1996 ASME Conference on Design Theory and Methods*, ASME 96-DETC/DTM-1520.

Science Information Resource Center, 1988,*Thesaurus of Scientific, Technical, and Engineering Terms*. Hemisphere Publishing Corporation.

TREC, (http://trec.nist.gov/)

Ullman, D.G., 1994, “Issues Critical to the Development of Design History, Design Rationale, and Design Intent Systems”, *Proceedings of Design Theory and Methodology - DTM ‘94*, DE-Vol. 68, pp. 249-258, ASME.

Wood, W. H. and Agogino, A. M., 1996, “A Case-based Conceptual Design Information Server for Concurrent Engineering”, *Journal of Computer Aided Design*, 28(5), pp. 361-369.

Wood, W.H. and Agogino, A.M., 1997, “A Prescription for Information Prospecting, Data Mining, and Design Refinement”, *Proceedings of DETC ‘97:1997 ASME Design Engineering Technical Conferences*, DETC97DTM-97-55

Yang, M. C. and Cutkosky, M. R., 1997, “Automated Indexing of Design Concepts for Information Management”, *Proceedings of the Eleventh Annual Conference on Engineering Design*, Tampere, Finland, August, 1997.

Yang, M. C., Wood, W. H., and Cutkosky, M. R., 1998, “Data Mining for Thesaurus Generation in Informal Design Information Retrieval” To appear in *Proceedings of 1998 International Congress on Computing in Civil Engineering*, American Society of Civil Engineers, Boston, MA, Oct. 1998.