Easing the Creation of a Multipurpose Case Library*

Eric Goldstein
eric@ils.nwu.edu

Smadar Kedar
kedar@ils.nwu.edu

Ray Bareiss
bareiss@ils.nwu.edu

The Institute for the Learning Sciences
Northwestern University
1890 Maple Avenue
Evanston, IL 60201

Abstract

As CBR case libraries grow, they will become use-ful for an increasing number of tasks. An open issue in CBR, therefore, is how to index cases so that they can be retrieved for multiple tasks. Currently, most CBR systems index cases using in-dexing schemes which are specifically tailored to their one task. One way current CBR systems can perform retrievals for multiple tasks is to use mul-tiple task-specific indexing schemes. A key draw-back to using multiple indexing schemes is the in-creased difficulty of the indexing task for the hu-man indexer. We present a method which makes it easier for a human indexer to create multiple task-specific indices. In particular, we describe how the method can (1) leverage off one kind of task-specific index to create another kind of task-specific index, and (2) partially automate the cre-ation of one kind of task-specific index. We will illustrate our method in the context of the Story Archive project at the Institute for the Learning Sciences.

Introduction

As CBR case libraries grow, they will become useful for an increasing number of tasks. An open issue in CBR, therefore, is how to index cases so that they can be retrieved for multiple tasks. The Story Archive project at the Institute for the Learning Sciences pro-vides a testbed for investigating the tools and tech-niques which will make possible the creation of a mul-tipurpose case library. The Story Archive is intended to become a large multimedia body of stories, including news footage, speeches, lectures, and interviews with

experts. Stories in the Archive will often be used for multiple purposes. For example, consider the follow-ing transcript from George Wallace's 1963 inaugura-tion speech as the democratic governor of Alabama:

"Today I have stood where Jefferson Davis stood, and took an oath to my people. It is very appropriate then that from this cradle of the Confederacy, this very heart-land of the great Anglo-Saxon Southland, that today we sound the drum for freedom.... In the name of the greatest people that have ever trod the earth, I draw the line in the dust and toss the gauntlet before the feet of tyranny. And I say segregation now! Segregation tomorrow! Segregation forever!"

The Story Archive currently retrieves this case for two different purposes. By *purpose*, we mean a task for which a case can be used. One task for the Story Archive is to explore the different sides of arguments in American society. The Story Archive attempts to both present an argument and a refutation of that ar-gument. After the user has viewed stories such as a documentary clip about Abraham Lincoln's views on states' rights or old news footage on Martin Luther King's views about integration, the Story Archive can attempt to refute their arguments about states' rights and segregation by retrieving the Wallace case. The Story Archive can also be used to put historical fig-ures into perspective, and give insights into seemingly unusual beliefs and behaviors. Wallace's view on seg-regation may have been typical for a white Southerner, but, when viewed as a democrat, Wallace was unusual in his opposition to racial integration. After the user has viewed a story about someone who has bucked the party line, the Story Archive can retrieve the Wallace case in order to put this unusual behavior into perspec-tive.

As we collect more and more cases, we need methods to ease the task of indexing the cases for multiple pur-poses. To develop methods which scale up, we must first decide how cases will be indexed, and then de-velop tools which make the cases easy to index. The issue of how to index cases for multiple purposes re-mains unresolved. It is widely acknowledged in the

CBR community that indexing schemes for case-based reasoners should be designed to fit the task for which the cases will be used [e.g. Hammond, 1989; Kolodner, 1989; Kedar-Cabelli, 1985]. This raises the following problem: if a set of cases are indexed with an indexing scheme which is tailored to a particular task, the cases will not be available for retrieval for other tasks. In general, using a task-specific indexing scheme makes it difficult to create case libraries which can be used for multiple purposes.

One approach to the problem of labeling a case for multiple purposes is to create a single indexing scheme which would provide sufficient structure and vocabulary to support all of the tasks for which the case might be retrieved. Initial work on a universal indexing scheme has been attempted [Schank et al. 1990; Schank and Fano, 1992]. Earlier versions of the Story Archive also used general-purpose indexing schemes [Bareiss et al. 1991; Schank et al. 1992]. Although research into general-purpose indexing schemes continues, we have found that it is hard to anticipate the requirements of an indexing scheme a priori.

Another approach to this problem is to label each case using multiple indexing schemes, one for each class of tasks. In the absence of an all-purpose indexing scheme, we are currently investigating the use of multiple task-specific indices.

A key drawback to using multiple indexing schemes is that creating multiple indices for a single case can be very time consuming. Even creating one index per case is hard work, and the need to create more than one threatens to make the job of indexing much more onerous. Specifically, an indexer is faced with two problems:

1. **The Feature Selection Problem:** The problem of deciding which aspects of a case may be relevant for indexing.

2. **The Index Derivation Problem:** The problem of finding and using the vocabulary of a representation theory to create an index which captures the selected features of a case.

The difficulty of creating indices delays the addition of new stories. For this reason, the use of multiple indexing schemes has made the development of indexing tools especially important.

In this paper, we describe a tool which accompanies one particular indexing scheme (a scheme which represents unusual biographic profiles). The tool eases the difficulty of creating multiple indices in the following ways:

1. **Feature Selection:** The tool helps the indexer select features from a case description by leveraging off indices already created for the case.

2. **Index Derivation:** The tool partially automates the task of index creation.

The tool's ability to leverage off already created indices has an effect which partially offsets the burden of using multiple indexing schemes: the greater the number of indices for a case, the easier it becomes to select features for a new index. In addition, by partially automating the creation of indices, the tool further reduces the burden of using the indexing scheme's representational vocabulary.

In the remainder of this paper, we provide background on the Story Archive project. We then describe our tool, illustrating it with a running example from the Story Archive. We conclude with a discussion of our ongoing research.

Background

In order to make the contents of the case library accessible and useful to users with different interests, the Story Archive is equipped with retrieval agents which retrieve stories for different purposes. Each agent exemplifies a different persona [cf. Laurel, 1990] with different retrieval goals. Given a story, the agents retrieve follow-up stories which serve their goals. In this paper we focus on two such agents, named Dickens and the Rabbi.

When the speaker in a story makes a particular point, the Rabbi has the goal of refuting the speaker's arguments by finding cases which illustrate counter-examples to the point made in the previous story. For example, given a video clip showing civil rights activist H. Rap. Brown calling for blacks to use all means at their disposal to achieve equality, the Rabbi can undercut Brown's point by retrieving a story showing Martin Luther King making a plea for non-violence.

The Dickens agent responds when the story makes it apparent that the speaker or main character has atypical traits or exhibits unexpected behavior. Dickens has the goal of giving the user insight into the main character's possible motivations and background by retrieving a story about a person with a similarly unusual biographical background. For example, given the case of a Democrat who advocates cutting government spending, Dickens can help explain why an American politician might buck the party line by retrieving the case of a republican who advocates increasing taxes.

Our tool helps create indices for Dickens. It does this, in part, by leveraging off indices created for the Rabbi. Both agents retrieve cases from the same domains, but use the cases for different purposes. To support these differences and commonalities, each agent has its own task-specific abstract indexing terms (which together make up what we call an indexing scheme), but each agent also has access to a commonly-shared *hierarchy of representational vocabulary* which currently represents the Story Archive's domain knowledge. The representational hierarchy contains both attributes and values, termed here as *categories* and *descriptive values*. For example, one category in the hierarchy is called "attitude-on-government-spending" and this category includes a range of descriptive values starting at "advocates-government-spending" contin-

uing through "de-emphasizes-government-spending" and ending with "rejects-government-spending."

We will now describe each agent's indexing scheme. The Rabbi uses a *point-based* indexing scheme which represents a set of descriptive values as argumentation points. The points of a speaker's argument are mapped out in terms of an abstract characterization of the speaker's argumentative strategy, and in terms of an intentional vocabulary of plans, goals, themes, and resources. For example, the Rabbi represents the speech by King mentioned above as a case of someone who advocates the plan of "non-violence" to achieve the goal of "racial-equality". Point-based indexing is described more fully in [Shafto, Bareiss, and Birnbaum 92].

Dickens uses indices which are based on anomaly. *Anomaly-based indexing* is a common technique in CBR [e.g. Schank 1982; Kolodner 1984; Hammond 1986; Kass, Owens, and Leake 1986]. Dickens uses the hierarchy of representational vocabulary in a more complex way than the Rabbi. For Dickens, the hierarchy also includes *stereotypes*. Each descriptive value is associated with a set of stereotypical attributes and their accompanying norms. For example, the descriptive value "democrat" has a large number of stereotypes associated with it. Dickens expects that a democrat will normally "advocate-government-spending." Dickens' indices represent expectation violations, in which person can be described by descriptive values which are atypical of the norm. For example, an index for Dickens would represent a democrat who atypically "rejects-government-spending."

Dickens' task-specific indexing scheme represents an anomaly by relating the anomalous value and the expectation violation with an *anomaly characterization*. For example, the cost-cutting Democrat would be represented as "a person who has a belief about government spending which is opposite what would be expected for a Democrat." In Dickens' representation, the relationship between the anomalous value and the norm is characterized as *opposite*. This relation is derived from the fact that the norm for a democrat, "advocates-government-spending", and the anomalous value, "advocates-spending-cuts", are at the *opposite* ends of the spectrum along the range of possible values in the category "attitudes-about-government-spending". Possible relationships between a norm and an anomalous value along the range of possible values are: *opposite, higher, and lower.*

An Indexing Scenario

We now describe an indexing session which illustrates our tool. We first show how an indexer creates a point-based index for the Rabbi. We then show how our tool can be used to ease the creation of anomaly-based indices for Dickens.

The session begins as a human indexer, choosing material for the Story Archive, decides to include material from a documentary on Wallace's 1963 campaign as the Democratic nominee for Governor of Alabama. The indexer decides to include a clip showing both Wallace's speech and a commentator's voice-over which explains that Wallace believes that Alabama must remain segregated because the federal government's plans for integration are an infringement on Alabama's states rights.

Creating a Point-Based Index

The human indexer engages in a multi-step process in order to create a point-based index for Wallace's speech. First, the indexer analyses the content of the speech, and tries to construe the speech in the terms of the system's representational vocabulary. Because Wallace's speech was political, the indexer consults a part of the hierarchy of representational vocabulary which is dedicated to political beliefs (see Fig. 1). The indexer scans down the hierarchy looking for descriptions which match the concepts in the video. The indexer finds the term "advocates-segregation", and then, looking for a concept concerning states' rights, the indexer finds the term "advocates-a-weak-union", a close match.

Next, in order to create the point-based index, the indexer casts these concepts in terms of an intentional vocabulary of themes, plans, goals, and resources. Wallace can be said to be advocating the plan of segregation in order to achieve the goal of a weak union. Finally, the indexer selects argument strategies from a set of *rhetorical templates*. Rhetorical templates are schemata for argumentative strategies which are based the way plans, goals, themes, and resources are used in an argument. From a hierarchy of rhetorical templates, the indexer chooses a template which says "plan is good because it achieves a good goal." The indexer instantiates this template to say that Wallace is advocating the plan segregation, because it achieves a good goal, a weak union.

By combining all of this information, an index for Wallace's speech is finally formed. This information is summed up in the following point-based index:

```
Clip Name: ''Segregation Forever"
Speaker:    George Wallace

Supports: ''Segregation is a good plan"
   Plan:            Advocates-Segregation
   Plan-Agent:      Americans
   Goal:            Advocates-A-Weak-Union
   Template:        Plan is good because it
                       achieves a good goal
```

Recall that the Rabbi's goal is to retrieve stories which oppose the point of the current case. Once the point-based index has been created, the Rabbi can use it to retrieve stories which oppose Wallace's point. The Rabbi retrieves a story about Abraham Lincoln to refute the notion that a weak union is a good goal, and it retrieves a video clip of a speech by Martin Luther

```
POLITICAL-BELIEFS

    CHALLENGE-ORTHODOXY?
        Advocates-Challenging-Orthodoxy
        De-emphasizes-Challenging
            -Orthodoxy
        Advocates-Orthodox-Position

    ATTITUDE-ON-DIVERSITY?
        Advocates-Uniformity
        De-emphasizes-Diversity
        Advocates-Diversity

    ATTITUDE-ON-TOLERANCE?
        Advocates-Intolerance
        De-emphasizes-Tolerance
        Advocates-Tolerance

    ATTITUDE-ON-EQUALITY?
        Advocates-Inequality
        De-emphasizes-Equality
        Advocates-Equality

    SOCIAL-INTEGRATION?
        Advocates-Integration
        De-emphasizes-Integration
        Advocates-Segregation
    ...

    GOVT-SPENDING
        Advocates-Govt-Spending
        De-emphasizes-Govt-Spending
        Rejects-Govt-Spending

    PERSONAL-WEALTH-ACQUISITION
        Advocates-Personal-Wealth-Acquisition
        De-emphasizes-Personal-Wealth
            -Acquisition
        Rejects-Personal-Wealth-Acquisition

    ...

    GOVERNMENTAL-UNION?
        Advocates-A-Union
            Advocates-A-Strong-Union
            Advocates-A-Weak-Union
        De-emphasizes-Issue-of-Union
        Rejects-Union

    MIGHT-MAKES-RIGHT?
        Advocates-The-Supremacy-Of-Force
        Rejects-The-Supremacy-Of-Force
```

Figure 1: A portion of the hierarchy of representational vocabulary used by both agents.

King to refute the notion that segregation is a good plan under any circumstances.

Discussion

A point-based index cannot be created quickly and easily. In creating the point-based index, the indexer had to consult a hierarchy of political concepts. The hierarchy we have created is quite lengthy, yet, as we are describing a work in progress, the hierarchy is far from complete. As our hierarchy grows larger, human indexers will face the *feature selection problem* as they comb the hierarchy looking for appropriate terms. Our ability to easily create multiple indices will be limited by the time it takes human indexers to examine representational hierarchies.

Creating a point-based index also takes a significant amount of abstract reasoning. After the indexer selected the most descriptive features, the indexer had to decide how to cast the selected features in terms of an intentional vocabulary of plans and goals. Then, again listening to the speech, the indexer had to think about the argumentative strategy Wallace used to connect his plans and goals. Manipulating this task-specific abstract vocabulary takes a great deal of inference on the part of the indexer. In general, in any system which uses retrieval-related abstract vocabulary in its indices, human indexers will face the *index derivation problem*. Our ability to easily create multiple indices will also be limited by the cognitive load incurred by requiring human indexers to manipulate task-specific abstract vocabulary.

Easing the Creation of Anomaly-Based Indices

After finishing creating indices for the Rabbi, the indexer uses our tool to create indices for Dickens. Recall that Dickens' indices describe the different ways a person in a story might be considered anomalous. The tool creates Dickens' indices in a two step process. First, the tool helps the indexer incrementally build a *profile* of the person in the story. A profile takes the form of a list of descriptive values. Second, after a descriptive value is added to the profile, the tool detects anomalies by comparing each descriptive value to the norm-based expectations associated with the other descriptive values in the profile. When an anomaly is detected, the tool creates an anomaly-based index.

When the indexer starts up our tool to create indices to describe Wallace (as he is presented in this case), the tool begins to help the indexer create a profile of Wallace by retrieving the point-based index described above. By looking in the appropriate slots, the tool finds two descriptive values: "advocates-segregation", and "advocates-a-weak-union." The tool uses these descriptive values as a jumping off point; the descriptive values become the initial elements of Wallace's profile.

The indexer is now ready to add to Wallace's profile by choosing new descriptive values from the hierarchy

of representational vocabulary. Since the indexer last used the hierarchy to select descriptive values which were political in nature, the tool directs the user's attention back to the politically oriented section of the hierarchy. The documentary clip describes Wallace as a democratic governor who had just won the his bid for re-election. The indexer selects "democrat" from the "political ideologies" part of the hierarchy, and "state-governor" from the "political occupations" part of the hierarchy.[1] (The indexer looks for attributes describing the ideas of "candidate" and "winner", but does not find them due to the current incompleteness of the hierarchy.)

To detect anomalies, each descriptive value in Wallace's profile is checked against the norm-based expectations associated with each of the other descriptive values in the profile (see [Leake 1989] for a similar but more elaborate approach to anomaly detection). Shown below are some of the expectations Dickens associates with "democrat". When the indexer adds "democrat" to Wallace's profile, the tool finds an incongruity: the profile includes "advocates-segregation", but Dickens' expectations for a democrat include "advocates-integration."

```
Democrat
    Reform?                  Advocates-Reform
    Social-integration?      Advocates-Integration
    Diversity?               Advocates-Diversity
    Tolerance?               Advocates-Tolerance
    Govt-Spending?           Advocates-Govt-Spending
    Govt-Taxation?           Advocates-Increased-Taxation
    Gun-Control?             Advocates-Gun-Control
    Abortion-Rights?         Advocates-Abortion-Rights
    Defense-Posture?         De-emphasizes-Defense
    Industrial-Policy?       Advocates-Industrial-Policy
    Personal-Wealth-         De-emphasizes-Personal-
      Acquisition?             Wealth-Acquisition
    Obedience?               De-emphasizes-Obedience
    ...
```

The tool then automatically creates an index which reflects the fact that George Wallace's support for segregation is anomalous for a democrat. To do this, the tool must first characterize the relationship between the expected value and the anomalous value. In the hierarchy of representational vocabulary, the category of "social-integration?" has three possible values organized into a range of attitudes: "advocates-integration", "de-emphasizes-integration", and "advocates-segregation". Because the expected value for democrat is at one end of the range, and the anomalous value is at the other, the tool recognizes that the two values oppose each other, and it characterizes their relationship as *opposite*. The following index, created by the tool, states that Wallace

[1]The political occupations part of the hierarchy has been adapted from *Scholl's Humanitome* [Scholl 1936], a writer's aid containing over two thousand categories for describing human activities and attributes.

can be viewed as someone who has a belief about integration which is opposite what one would expect of a democrat.

```
View:  Democrat
    Spec-of:    Politically-Left-Of-Center
    Category:   Political-Ideology

    Expectations
        Social-Integration?
            Norm:  Advocates-Integration
            Opposite:  "George Wallace", ...
    ...
```

The tool does not only use Dickens' expectations to detect anomalies; it also uses them to offer suggestions. If the indexer requests assistance in creating a profile, the tool can present the expectations it associates with each descriptive value in the current profile. Each expectation is suggested as an addition to the profile. For each suggested value, the indexer is given three choices: a suggestion can be ignored as irrelevant to the actual case, it can be added to the profile, or it can be marked incompatible with the actual case. If the indexer indicates that one of the expectations is incompatible, the indexer is given the opportunity to select an alternative value from the same category. The selected alternative value is then added to the profile. Because the tool's suggestions are derived from expectations, the selected alternative value will always be incompatible with another value in the profile, and the tool will automatically create a new anomaly-based index to reflect this incongruity.

In the case of George Wallace, when the indexer asks for suggestions, the tool presents Dickens' expectations for each of the descriptive values in the profile: "advocates-segregation", "advocates-a-weak-union", "Democrat", and "state-governor." The indexer scans the list of expectations, looking for values to select. One of the expectations associated with "advocates-segregation" is "advocates-inequality". Because the indexer finds it interesting that Wallace did not, in fact, advocate inequality, the indexer indicates disagreement with this suggestion and is presented with the range of values from the category "support-for-equality." Recall that the values for this category are organized into a range from "advocates-inequality" through "de-emphasizes-equality" to "advocates-equality". The indexer chooses "de-emphasizes-equality", and the tool creates a new index based on this choice. In the hierarchy of representational vocabulary, "de-emphasizes-equality" is farther along the range of values than the expected value, "advocates-inequality", and so the the tool characterizes the relationship between the alternate selection and the expected value as *higher*. The following index, created by the tool, states that Wallace can be viewed as someone who has a level of support for equality which is higher than what one would expect of someone who advocates segregation.

```
View:  Advocates-Segregation
  Category:  Support-for-Segregation

  Expectations
    Support-for-Equality?
        Norm:  Advocates-Inequality
        Higher:  "George Wallace", ...
    ...
```

Once these indices have been created, Dickens can retrieve stories about people who are analogous to Wallace. For example, using the first index, Dickens analogizes the Wallace case to a story about Lestor Maddox, another democratic governor who supported segregation. Using the second index, Dickens analogizes the Wallace case to a story about the early career of Malcolm X, another person who advocated a form of segregation and yet did not advocate inequality. Dickens can also offer more distant analogies by generalizing the elements which make up the index. For example, returning to the first index, if Dickens continues to focus on unexpected beliefs about segregation, but generalizes the index from "Democrat" to "politically-left-of-center," Dickens can retrieve stories about other people who had views about segregation which were opposite what would be expected from someone who was left of center. (And if we index Malcolm X as someone who was left of center, we find that Malcolm X can again be analogized to Wallace, thereby supporting our earlier analogy.)

Discussion

Our tool partially alleviates both the feature selection problem and the index derivation problem. The tool eases the feature selection problem by leveraging off previously created indices. Information from previously created indices is used in the following three ways.

First, the tool directs the users attention to the parts of the hierarchy of representational vocabulary which were useful in previous indices. This simple method is helpful when a descriptive value describing the content of the person's speech is similar to an unusual aspect of the person. In the example case, since Wallace had already been described in terms of his political beliefs, the tool directed the user to the "political" part of the hierarchy. Because Wallace had an unusual view with respect to his political ideology, it proved helpful to direct the user to this part of the hierarchy while selecting features.

Second, the tool uses information from previously created indices by adding descriptive values from the old indices to the profile. The descriptive values added to the profile are used to generate additional expectations, and sometimes these expectations are incongruent with other descriptive values in the profile. When this happens, the information from a previously created index is directly incorporated into an index. In the example case, the descriptive value "advocates-segregation" was imported from a previously created index and became the basis for the first anomaly-based index created for Wallace.

The third, and by far, the most consistently helpful way the tool uses information imported from previously created indices, is to make suggestions to the user. The additional expectations generated from the imported descriptive values are presented to the user as suggestions. The tool's ability to present expectations to the user as suggestions is quite helpful in alleviating the feature selection problem. By presenting the suggestions, the tool directs the indexer's attention to the descriptive values which are most likely to be both relevant to the case, and incongruous with other descriptive values in the profile. As the indexer adds more descriptive values to the profile however, the list of suggested expectations becomes longer in turn, and an increasing number of the suggestions will be irrelevant to the case at hand. Nonetheless, the list of suggestions will normally be far smaller than the entire representational hierarchy, and the contents of the list will often have a far greater density of relevant values.

Once the indexer selects a set of features, the tool eases the index derivation problem by performing all additional inference. The indexer did not have to decide which combinations of descriptive values were unusual, nor did the indexer have to use abstract vocabulary such as *opposite* or *higher*. Because the tool detects anomalies and characterizes the relationship between the norm-based expectation and the anomalous value, the indexer only needed to select descriptive values to create an index.

Future Work and Conclusion

Dickens and the Rabbi both use their own indexing schemes, and share the same representational hierarchy of domain knowledge. In this paper, our examples have shown how domain knowledge can be mapped over from a point-based index to an anomaly-based index. We have further shown how our tool can eliminate the need for the human indexer to use Dickens' task-specific abstract indexing scheme. It would be ideal, of course, if we could propose a more general method. Such a method would go something like this: To ease the creation of a multipurpose case library, (1) use task-specific indexing schemes which share domain knowledge, (2) develop methods to map over domain knowledge, (3) automate the index derivation process in order to isolate the indexer from the abstract task-specific vocabulary which facilitates retrieval.

There are several issues which must be examined before we can seriously propose such a method. First, we have only examined two task-specific indexing schemes thus far. We need to investigate and categorize other indexing schemes. Second, it seems reasonable to suspect that many tasks will have little or no domain-dependent knowledge in common. It is not yet known

17

whether abstract thematic knowledge is enough to bridge the gap between disparate tasks, and this issue must be further explored. Finally, it is not clear that automatic index derivation is possible for all indexing schemes. It is this last point that we will further address.

Our tool is able to derive indices for Dickens given only two descriptive values from the indexer. It seems clear that this trick will not work for point-based indexing. Without extensive world knowledge, there is no way that a tool given two descriptive values could infer that, for example, a speaker was advocating one of the descriptive value because they believed that another descriptive value was good. Our tool also had access to previously created indices, but again, it seems unlikely that additional indices would make it possible for a tool to automatically derive a point-based index. Until recently, it appeared that our method might not extend beyond anomaly-based indexing schemes.

We now believe the method may be more extensible. We have recently begun to investigate the possibility of interleaving the process of indexing stories with the process of using the Story Archive for its intended purposes. We[2] are currently designing a tool to automatically create point-based indices within the context of using the Rabbi. An indexer would hear a story, see its point, disagree with the system, and then offer a new story to refute the previous point. By using the Rabbi's knowledge of argument strategies, it may be possible infer a new index (thus easing both the feature selection problem and the index derivation problem). We are investigating ways in which an anomaly-based index could contribute to the process of index derivation (one approach is to refute claims by finding one anomalous exception). If these approaches can be successfully developed, it would bolster our general claim that previously created indices can be combined with index derivation methods to assist in the creation of a multipurpose case library. In any case, this paper represents one step towards multipurpose case-based reasoning systems.

Acknowledgments

This paper was spurred by a discussion with Marvin Minsky. We want to thank Eric Shafto for his explanation of point-based indexing. Thanks to Dick Osgood, Tom Hinrichs, Chris Wisdo, and other ILS colleagues for their comments on drafts of this paper.

References

Bareiss, R., Ferguson, W., and Fano, A. [1991] The Story Archive: A memory for case-based tutoring. In *Proceedings of the 1991 DARPA Case-Based Reasoning Workshop* (pp. 267-279). San Mateo, CA: Morgan Kaufman.

Hammond, K. [1989]. On functionally motivated vocabularies: An apologia. In *Proceedings of the 1989 Workshop on Case-Based Reasoning*, pages 52-56. Morgan Kaufmann.

Kass, A. M., Leake, D. B., and Owens, C. C. [1986]. Swale: A program that explains. In *Explanation Patterns: Understanding Mechanically and Creatively*, pages 232-254. Lawrence Erlbaum Associates, Hillsdale, NJ.

Kedar-Cabelli, S. T. [1985] Purpose-Directed Analogy In *Proceedings of the Cognitive Science Society Conference, pages 150-159*.

Kolodner, J. [1989]. Judging which is the "best" case for a case-based reasoner. In *Proceedings of the 1989 Workshop on Case-Based Reasoning*, pages 77-81. Morgan Kaufmann.

Kolodner, J. [1984]. *Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Laurel, B. [1990]. Interface Agents: Metaphors With Character. In B. Laurel (Ed.), *The art of human-computer design*, pages 355-366. Reading, MA: Addison-Wesley.

Leake, D. [1989] Anomaly Detection Strategies for Schema-based Story Understanding. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*, pages 490-497. Ann Arbor, MI. Cognitive Science Society.

Schank, R. and Fano, A. [1992]. A Thematic Hierarchy For Indexing Stories In Social Domains. Technical Report 29, The Institute for the Learning Sciences, Northwestern University, Evanston IL.

Schank, R. et al. [1992]. Agents in the Story Archive. Technical Report 27, The Institute for the Learning Sciences, Northwestern University, Evanston IL.

Schank, R. et al. [1990]. Towards a general content theory of indices. In *AAAI Spring Symposium on Case-Based Reasoning*, pages 36-40, Palo Alto, CA. American Association for Artificial Intelligence.

Schank, R. [1982]. *Dynamic Memory*. Cambridge University Press, Cambridge, England.

Shafto, E., Bareiss, R., and Birnbaum, L. [1992] A Memory Architecture for Case-Based Argumentation In *The Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society, pages 307-312*.

Scholl, C. V. [1936]. *Scholl's Humanitome*. Verbis Publishing Co. New York.

---

[2]This idea is being pursued in conjunction with Eric Shafto.