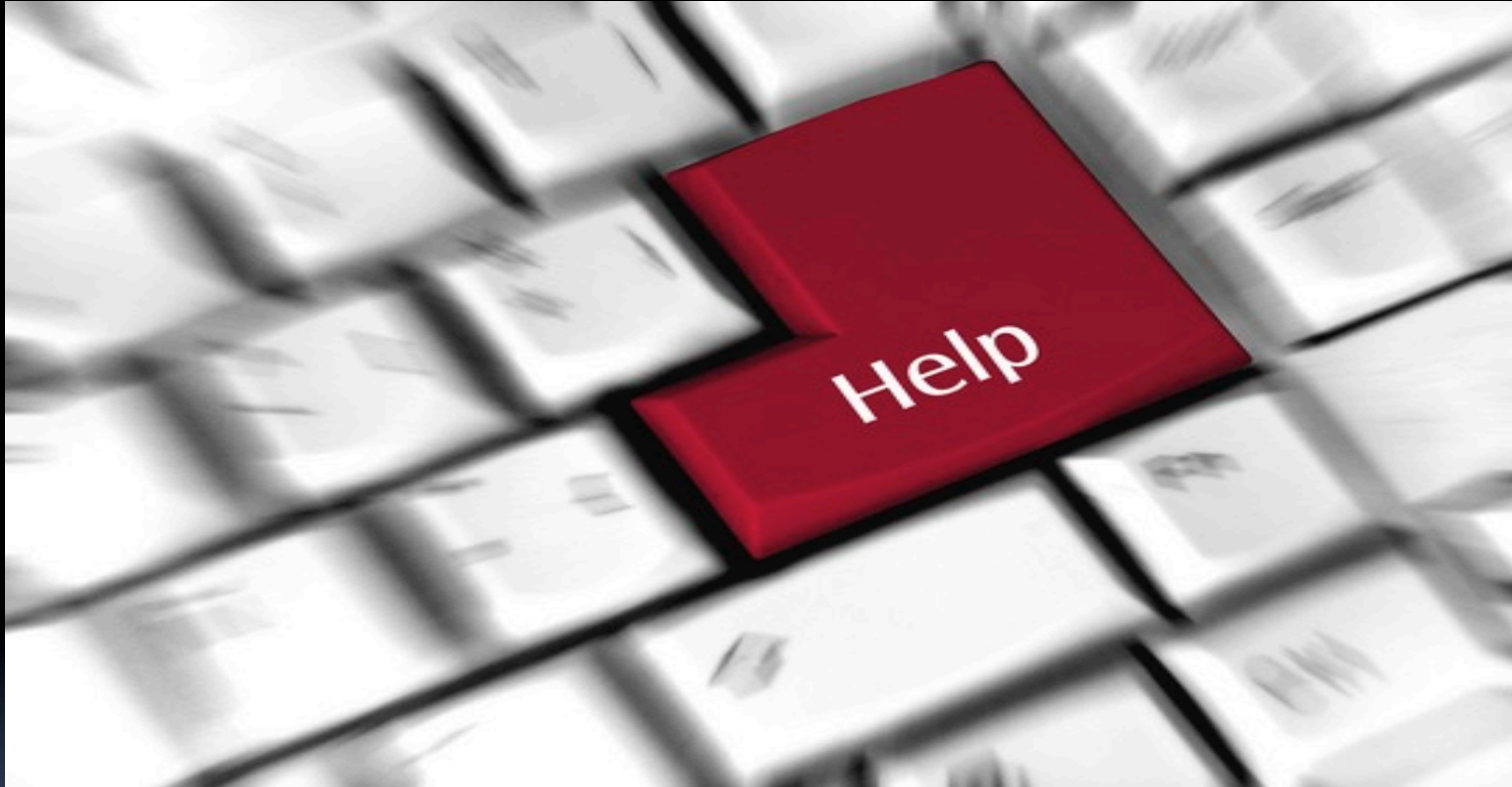


HELP DESK SYSTEMS

Using “CaseBased Reasoning”



K. SIDDARTH YAGNAM
CSE-435(FALL 2013)
LEHIGH UNIVERSITY

Topics Covered Today

- What is Help-Desk ?
- Components of HelpDesk Systems
- Types Of HelpDesk Systems Used
- Need for CBR in HelpDesk Systems
- GE Helpdesk using ReMind System (1993)
- CBR-TM
- The HOMER System (Will Not be Discussed Today)
- References

What is HelpDesk?

- A **'help-desk'** is an **information and assistance resource that troubleshoots problems**
- Many corporations provide helpdesk support to their customers

Companies Using CBR:

- ✧ AT&T Bell
- ✧ UK Department of Social Security
- ✧ Honeywell
- ✧ British Telecom
- ✧ Mercedes-Benz




Components of HelpDesk

- User
- Desk-side Team
 - ✧ Desk consultants
 - ✧ Phone Consultants
- Network Team
- Server Team
- Other Team
 - ✧ Library, WIRED, etc



Types of HelpDesk Systems

- Text Based HelpDesk Systems
 - ✧ Text Search
 - ✧ Keyword Search
 - Maintains a Natural text data
 - Rule Based Systems
 - ✧ Maintains a special Knowledge Rule set
 - Case Based Reasoning Systems
 - ✧ Maintains all experience in the form of cases
- 

TextBased Helpdesk Systems

- Text or keyword search techniques access information from past cells by making directed queries
- In keyword Approach, past cells are annotated with Keywords
- The query for the new case also must be annotated with relevant keywords
 - ✧ Cases having more keywords in common are retrieved
- In text search , each case is stored with a free form text description of the problem
- The comparison is between the text description of the new case with all the past cases
 - ✧ Cases having more character or words substrings in common are retrieved

Advantages of Text-Based Systems

- The knowledge of both the systems grow automatically with the # of calls/problems received
- The Accuracy of both the text search and keyword search 'can' be very good (???)
 - ✧ GIVEN all the users do use the same “keywords” or the same kind of “text description” for all cases
 - ✧ The description must not be too over-complicated or too vague
- Text Systems need not even maintain a set of keywords associated with past cases



Disadvantages

(As expected)

- Accuracy is more dependent on consistency and diligence of the users (Which can be never expected)
- The larger the User community the wider the range of keywords, which effect the retrieval accuracy (Missing Cases with Similar but not exact keywords)
- Text search will be more messy, as we can describe the same problem in many different ways
ex: “Can’t Write/Read data from disk”
(or)
“Can’t access data on my Seagate-Go Flex”



Precision and Recall

- Recall measures how likely a given system returns the information you are looking for.
- Precision measures the amount of relevant information returned.
- In general, increasing the precision tends to reduce its recall efficiency and vice versa.
 - ✧ If you are too general, then you can't find the solution in all the data.
 - ✧ If you are too specific, the system offers no good solution at all.



Rule-Based Systems

- Distill the relevant information from individual call records into a structured set of 'Questions' that can be used to systematically detect and resolve any problem.
- Eliminates the need for user to guess what information will be required.
- This operates quickly as rules are organized as a hierarchical decision tree.
- Store only knowledge from previous cases but not the cases.

Disadvantages

- Collection (extracting and encoding) of Rules is Time consuming
- Very Difficult to update and maintain
- Useless in situations not planned for or not developed i.e. new kinds of problems
- Cant update once encountered (static)
 - ✧ Unless someone adds information to form new rules
- Overly restrictive
- Just prints out the best solution but cannot cite the actual or relevant past case (As they don't store them)

CBR in HelpDesk Systems

- The main intuition behind this is (we all know)
 - ✧ “If Symptoms of two problems are similar, their diagnosis is also similar”
 - ✧ Significant portion of human cognition and problem solving involves recalling entire prior experience, rather than just a piece of knowledge.
- By using CBR we can inherit all the features of the above methods and also overcome their drawbacks
 - ✧ We can add new knowledge just by adding a new case (as they are independent of one another)
 - ✧ We can overcome the problem of inconsistency from the above methods by organizing different type of information into a single coherent structure.
 - ✧ Using Machine learning we can automatically derive the relevant case features rather than building and maintaining a set of Rules



Building a Case-Based System

- We have data collected in the form of online logs
- Creating the Case-Base
 - ✧ Collecting the Data
 - ✧ Extracting features from Data
 - ✧ Indexing the Data
 - * Indexing scheme
 - * Retrieval scheme
- Testing the Case-Base



Using the above methods we do convert raw tickets into a Case-Base



Pre and Post Goals

- What do we have ?
 - ✧ A fair amount of logs from existing call tracking Database
- What do we need?
 - ✧ Transforming them to diagnostic Cases
 - ✧ And a case retriever
- Platform & System used?
 - ✧ ReMind
 - ✧ Macintosh



Collecting the Data

- GE provided several hundreds of recorded online work-logs
- For a good CBR we need a “well-distributed” set of historical cases
- Each record has an identification number, an “initial description of the problem” and operators recommendations or analysis

Call Tracking Log :

PROBLEM ID	DESCRIPTION	ANALYSIS/RECOMMENDATION
5731	User is reporting that "the mouse doesn't work well anymore"	DD Checked mouse cable: OK. Removed ball from mouse, cleaned ball and rollers. This solved the problem. Recommend user purchase mouse-mat. Will call in two weeks to see if problem is recurring.
5735	James says that Word for Windows can't print to his laserjet IIp	DD System did not have enough memory to start printing while Word was running. Turned on VM as quick fix, told him to submit requisition for more RAM.

Extracting features from data

➤ Text Features

- ✧ Significant part of representation is based on raw text supplied
- ✧ It had to be cleaned* (Ensure it will be trained on representative, clean data)
- ✧ Deleted those cases that are irrelevant to the Domain*
- ✧ Making minor changes to the names or standardizing the Product Names*

➤ Developed a new representation!

- ✧ Computationally inexpensive as keyword matching.
- ✧ But Captures meanings of keywords.

Steps In Developing

- Created a Hierarchical tree of general concepts and individual symbols*

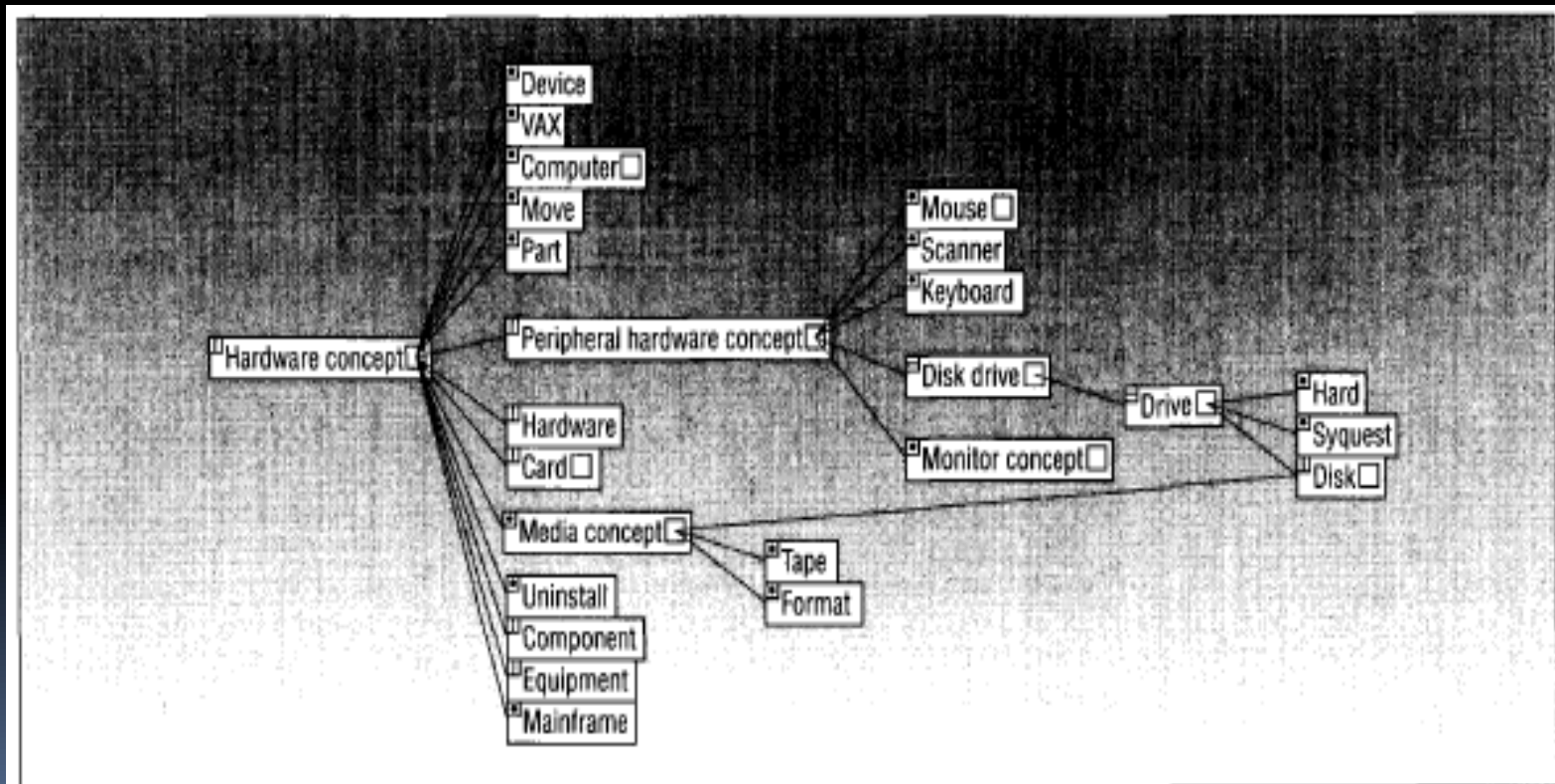


Figure 2. Part of the defined symbol hierarchy. The nodes containing squares are nonterminal.



Steps In Developing (continued)

- Using the Morphology filter & a Symbol lookup Function converted the data to a list-of-symbol fields

The system identifies the words in the hierarchy and collect them to a list.

“I cant save any documents on my floppy disk”

to

“can’t save document floppy disk”

Differences between this and Key-word Search

Depends on hierarchical structure, So words are inter-related and have synonyms – (This is a boon and a disadvantage)

Ex: “SyQuest cartridge” and hard drive.

- Multiple Inheritance in memory Hierarchy

Including New Features

- Some features which are missing in the original tasks (in the feature)
 - The Operating System
 - Hardware Platform
 - Software Application
 - List of problems (Rather than a single precise one)
- This creates a more precise and consistent description of a Case Base and ultimately better diagnosis of the problem (In General)



Demo of Our College HelpDesk System

“Footprints”

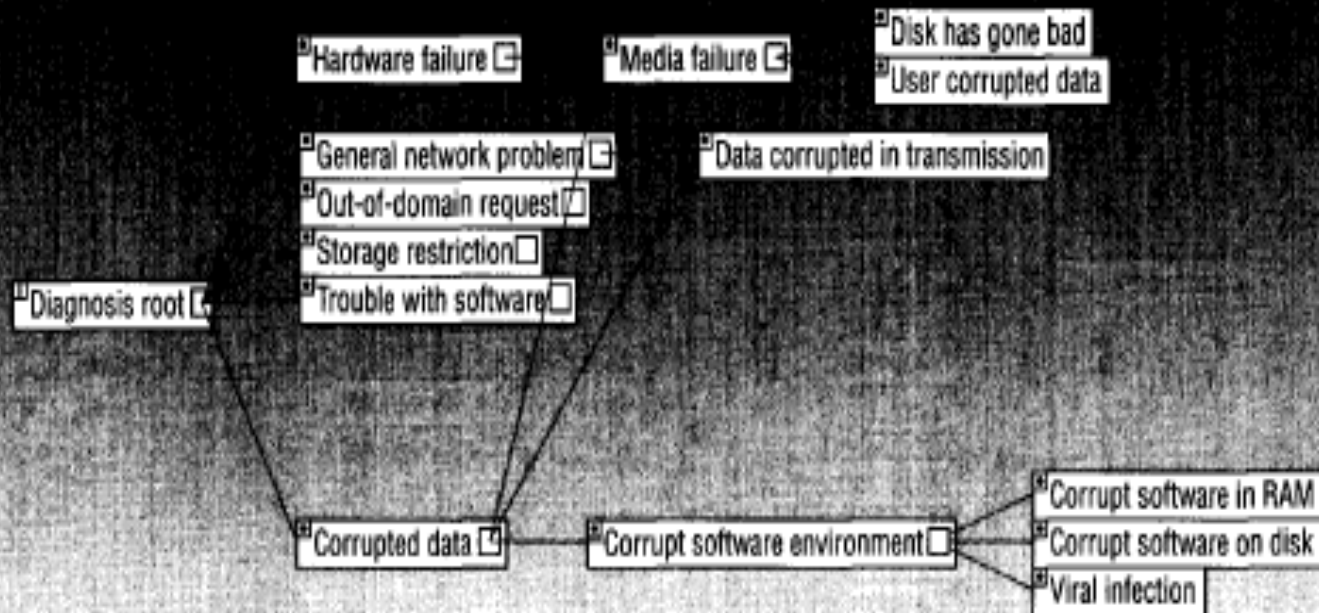
- UniPress(R) Software, Inc.

May-2005

Indexing the Data

- Two main kinds
 - That doesn't require cases to be categorized
 - KNN is an example
 - That require cases to be categorized
 - We Index a case based on both the features and also the diagnostic category associated with it
Ex: If we are unable to access a disk, there may be two possibilities, either the file system has gone bad or the disk has gone bad
- Once problems are divided, then an inductive learning algorithm had been used to create a classification tree

Some diagnosis Symbols



Hybrid Retrieval

- Two Phase Retrieval System
 - Inductive retrieval and
 - Nearest neighbor ranking
- The cases are scored and compared to the problem being diagnosed

$$\text{Total Score} = \frac{\sum(\text{similarity}(\text{In}[i], \text{Ret}[i]) * \text{Weight}[i])}{(\text{Totalweight})}$$

- Removes the “Duplicated” cases

Data Entry Windows

Operating system

Hardware platform

Peripheral device

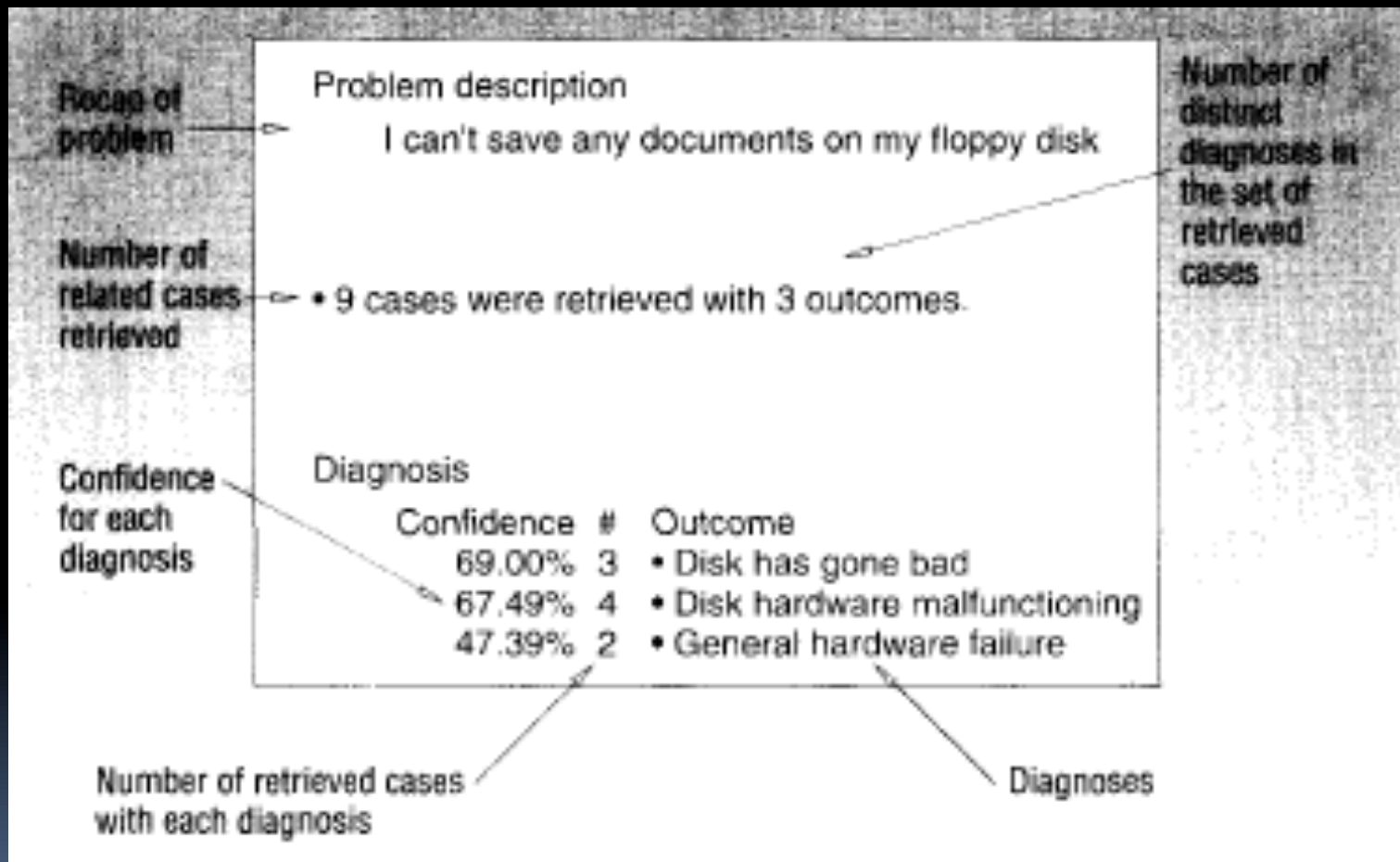
Application software

Software symptom list...

Problem description..

ReMind
for Help Desks

Result Window



Time Taken To build the System

- Total Prototype took 4 Man months
 - 2 Months to build the Library (200 cases)
 - 2 Months to construct the interface
- Deriving the Data from the, problem resolution section, is hard as it contains both solution & recommendations
- Important Part is Everything is done “Manually”
 - Constructing the hierarchy of the words
 - Cleaning, Deleting unnecessary cases, changing Names

Conclusion

- One of the First attempt to use CBR in Helpdesk systems
- Construction is mostly done manually
 - So time consuming
- Adding new cases is easy
- Not Scalable



Authors:

Juan Angel, Stella Heras Barbara , Vincent Botti

CBR-TM
CBR-TM

A New Case Based Reasoning System

A New Case Based Reasoning System

Introduction

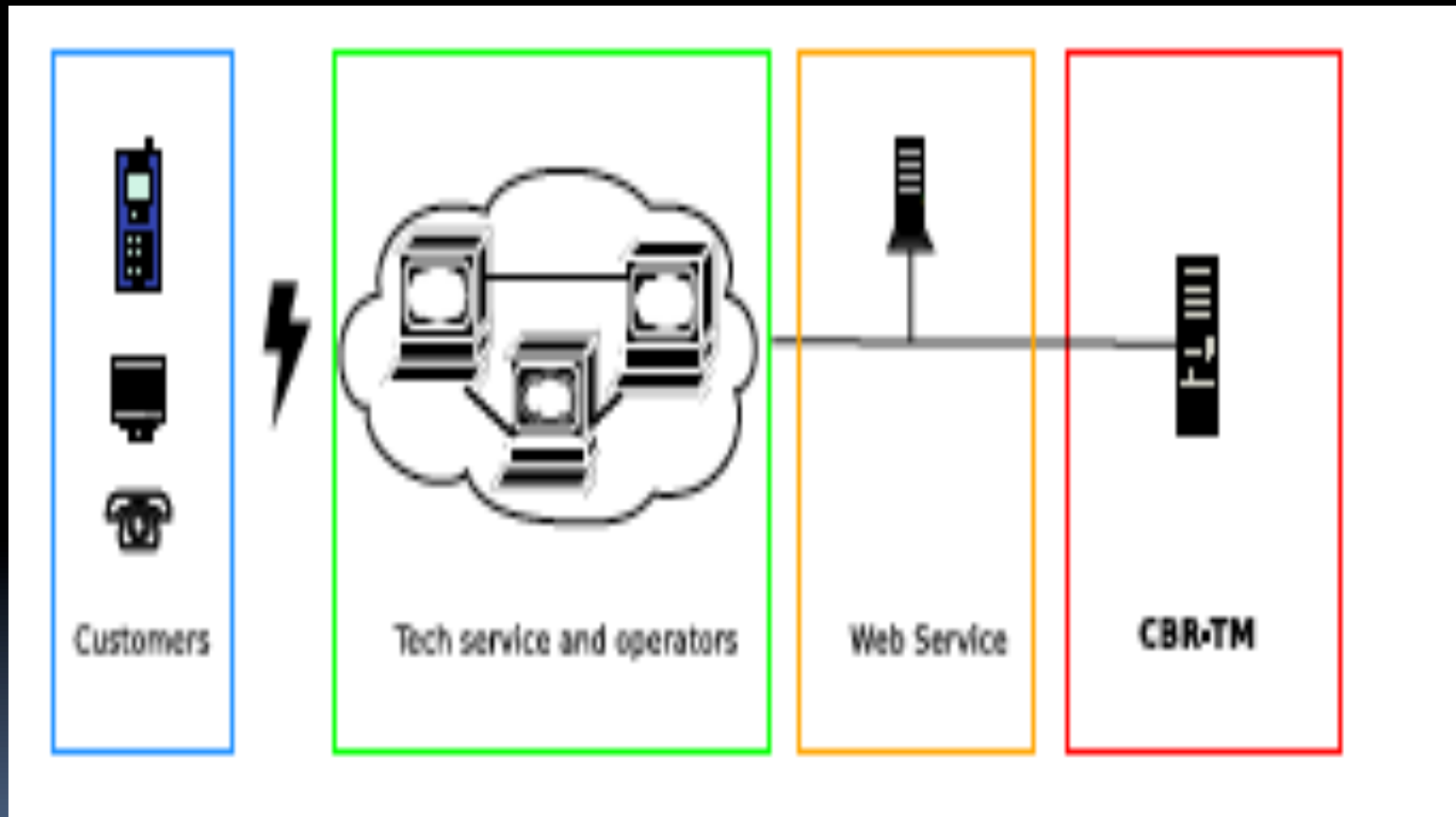
- An Intelligent module that works with “I2TM”
- A more generic CBR system
 - Provides Intelligent multi Domain support
- The system is independent of any change in the help-desk system
- Has been tested in a specific HelpDesk Environment that provides support for computer systems



I2TM

- Developed by a Spanish company “Tissat”
- Each request received is stored as a ticket
- These requests vary from computer problems to public services
- Main goal is to, rapidly solve these requests from a very diverse domains
- CBR-TM acts as an Intelligent module

Architecture Overview

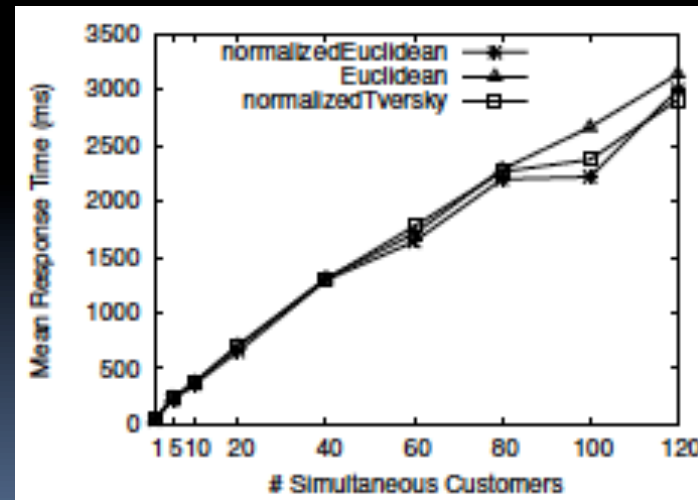
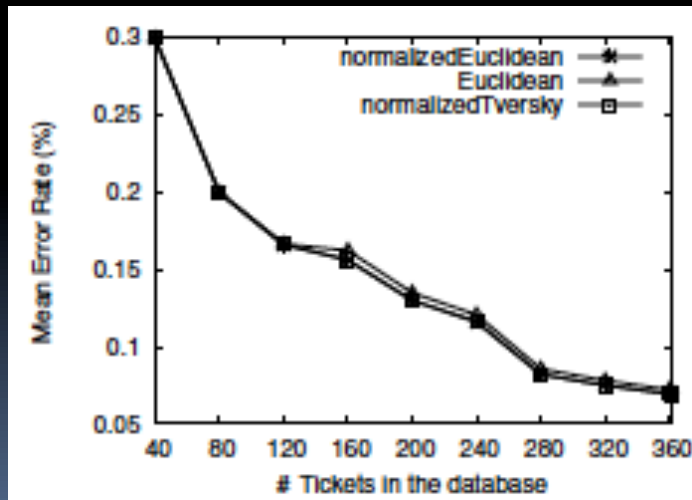


THE CBR-TM Module

- So each ticket consists of a set of categories that classifies it belongs to a certain type of problem
- A case is a prototyped representation on a set of tickets with same features and same successfully applied solutions
- The main challenge is, be able to work with heterogeneous tickets and must also be able to compute similarity between them
- Retrieval is made using Euclidian similarity measures
 - Normalized
 - Classic
- If there is no case similar enough to the new ticket, the system stores this ticket and its solution in the case base in the retention phase

Results

- Used a database of tickets* that came from the computer errors
- First they are trained using the tickets (loading the casebase), and then tested with the new tickets



Conclusion

- Although CBR-TM has been tested only in solving computer error, Tissat plans to use it in the different domains
- The learning is very effective
- Simultaneous requests are handled quickly (compared to most other systems)
- Main Future interest is to change the characterization to automatic



Authors:

Mehmet Göker , Thomas Roth, Wolfgang Wilke

HOMER System

HOMER SYSTEM





The Three Service Levels of a Help Desk System

- End User
- Hotline (Level 1)
- System Administrator, Application Specialists (Level 2)
- Maintenance and OEMs (Level 3)



Structure and Representation of a Help Desk System

- Attribute Value Pair
 - Good for answering trivial question
 - Good when users of the help desk system are inexperienced
- Object Oriented Representation
 - Structure of the technical system to be diagnosed can be represented in the necessary degree of detail
 - Symptoms can be clearly related to the object to which they belong to
 - The semantics of the problem description can be captured and used for selecting appropriate prior experience
 - A high retrieval accuracy can be achieved



Case Model (Problem)

- The Topic
 - The area in which the problem is located
 - hardware, software, network, printing service etc...
- The Subject
 - The physical object that the failure is related to
 - Specific software, printer, router, etc...
- The Behavior
 - The way the subject behaves
 - Wrong print size, screeching sound, no dial tone, etc...



Case Model (Situation)

- A set of attribute-value pairs describing symptoms that are important to diagnose the fault
 - Contain the minimum amount of info required to diagnose the problem (independence, completeness, minimalist)



Case Model (Solution)

- Contains the fault and the remedy
 - Composed of text or hypertext links
 - Can include links to more detail description
 - Can be a result of various situations
- Each complete path from problem to solution becomes its own case



Kinds of Cases

- Approved cases
- Opened cases (everyone can see)
- These case are separated into a case buffer (opened) and a main case base (approved)



User and Roles

- Help desk operator
 - Lowest access rights
 - Use application on a regular basis to solve problems
 - Case retrieval and case acquisition
- Experience author
 - Responsible for case maintenance and case approval
 - Checks for redundancy and consistency
- Experience base administrator
 - Creates and maintains the domain and case model
 - Administer users and access rights



Client/Server Architecture

- Allows all users to get the same up to date information
- Eases the maintenance of the domain model and the case base

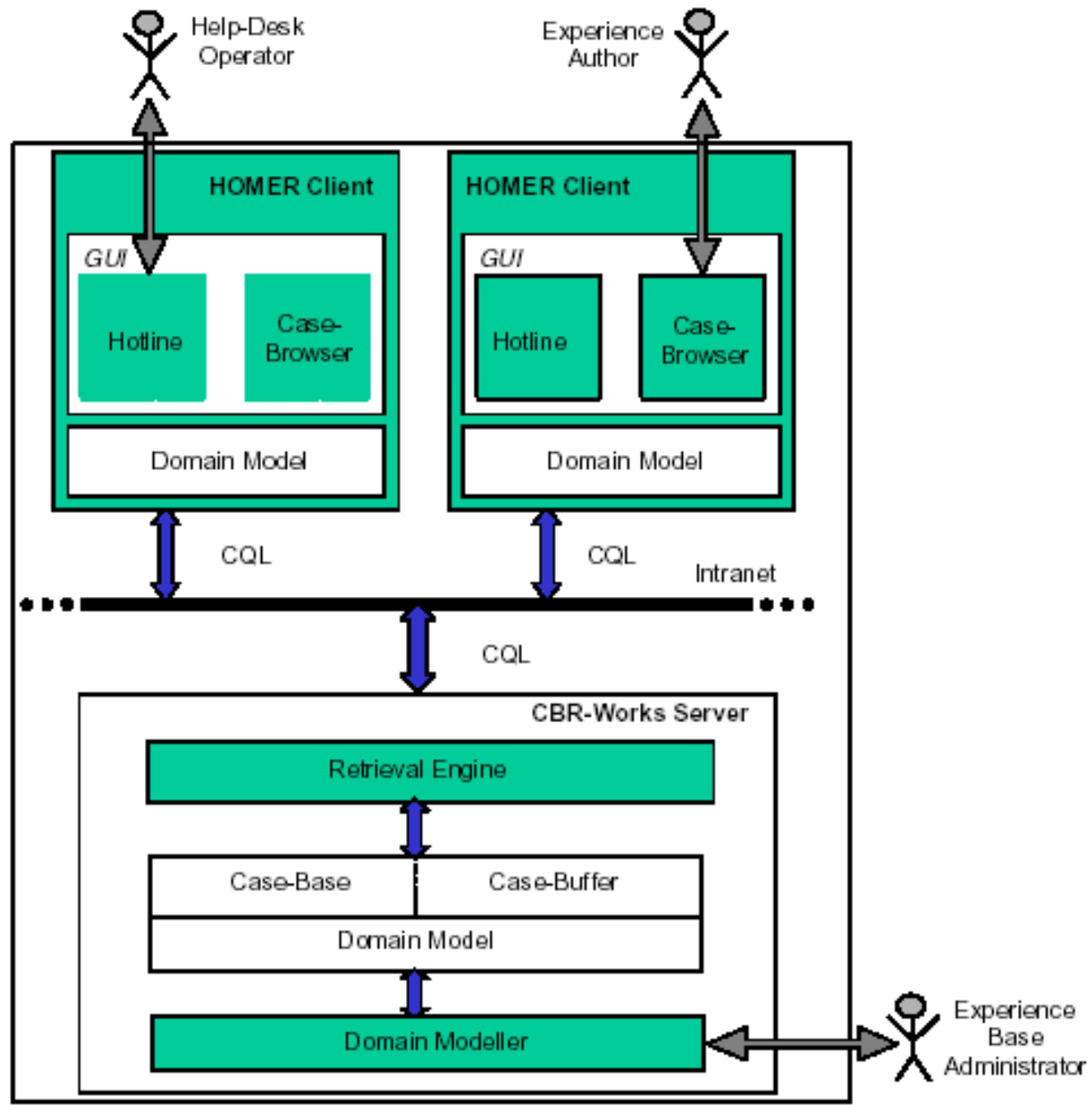


Fig. 11.3. The HOMER Architecture.



Retrieving Problem Solutions with Homer

- Two modes
 - Manual
 - User can enter as much information about a problem as wanted and then invokes a retrieval method
 - All matching case are retrieved
 - Automatic
 - The system retrieves matching cases after every item entered
- Solutions are displayed in the bottom view by decreasing relevance (CCBR)



Feedback

- Can be retained by pressing the retain button
 - Opens a case entry interface
 - Operator can make final modifications
 - Document why the case should be keep

Solutions

Lösungsvorschläge

Relevanz	Fall-Name	Lösung
0.941557	Druckproblem3	plt_client zum Au
0.941557	Plotproblem1	plt_client auf Ans
0.656557	CHelpDeskCase	Keine Lösung vo
0.651111	Plotproblem2	plt_serv auf Plot

Fig. 11.7. Suggested Solution.



Case Browser

- Used by the experience author to manage the case base
 - Case creation
 - Case copy
 - Delete case
 - Approve case

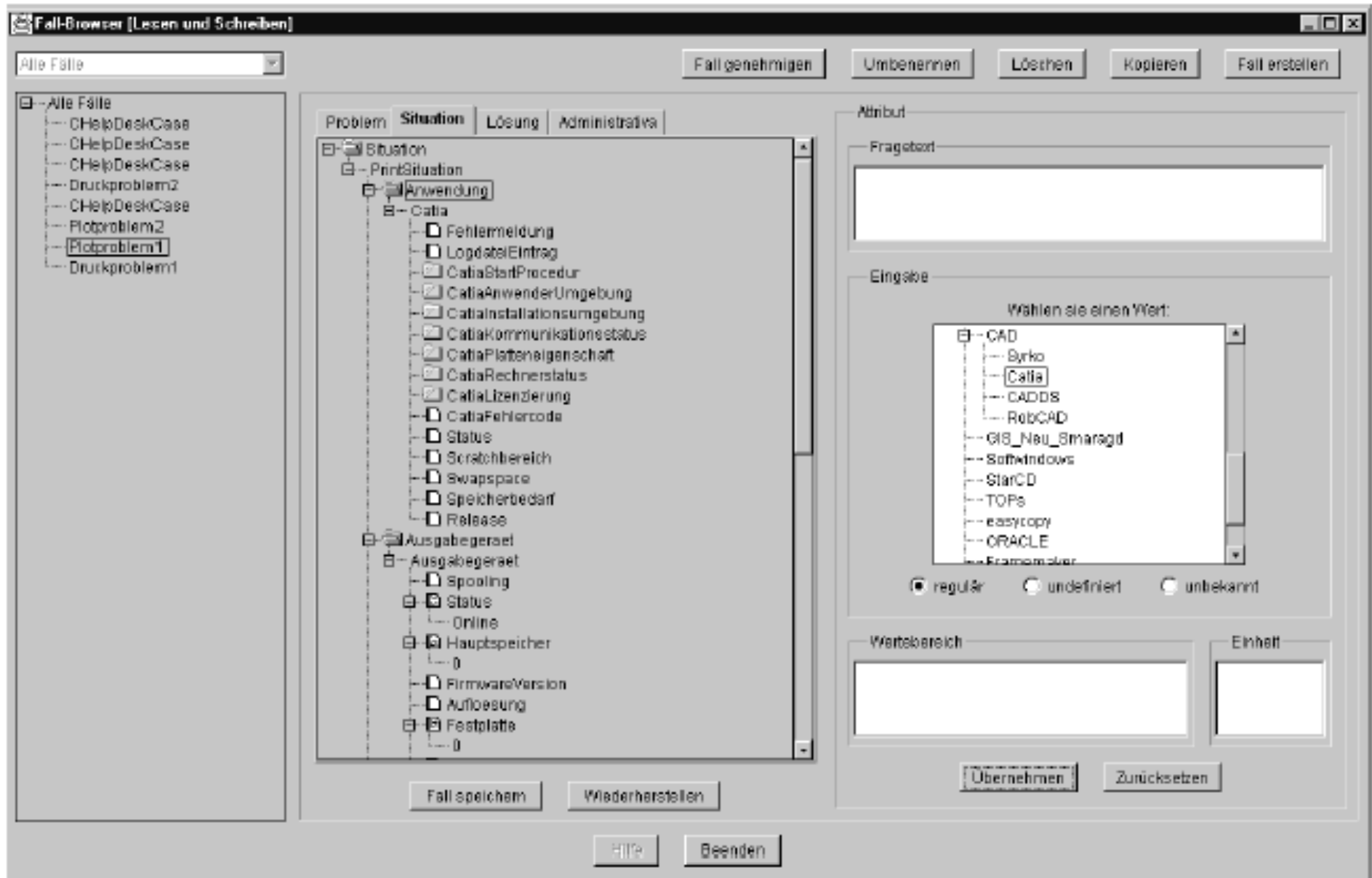


Fig. 11.8. The Case Browser.



The Development of the Homer System

- Initial knowledge acquisition
 - Three goals
 - Training the project team in knowledge acquisition
 - Initializing the knowledge in the system
 - Collecting enough help-desk cases

Overview of the Design and Maintenance of Help-Desk system

- Project planning and initialization
- Implementation of a rapid prototype
- Evaluation and revision of the prototype
- Implementation of the integrated case-based help-desk support system
- Evaluation and revision of the case-based help-desk support system
- Utilization of the case-based help-desk support systems



Evaluation of Homer

- Performed by INRECA II
 - Benefits for help desk operators
 - 102 problems of which 45 trivial or directed to the wrong help desk Homer solved 18 (32%)
 - Time to solve without = 141 min with = 9 min
 - Results better than expected



Summary

- Help desk systems
 - help solve problems faster
 - give more people more knowledge
- Building
 - long and difficult
 - need to convince people go give up their knowledge
- Maintaining
 - Requires constant maintenance



References

- 1 . Chapter 11. Experience Management for Self-Service and Help-Desk Support
- 2 . http://www.it.iitb.ac.in/~palwencha/ES/J_Papers/CBR_HelpDesk2.pdf
- 3 . https://www.haiti.cs.uni-potsdam.de/proceedings/ECAI-06/Proceedings/ecai/posters/ECAI06_199.pdf
- 4 . HOMER Systems (1999)

Thank You

