CBR Model for Freeway Work Zone Traffic Management

Asim Karim[1] and Hojjat Adeli, F.ASCE[2]

Abstract: A case-based reasoning (CBR) model is presented for freeway work zone traffic management. The model considers work zone layout, traffic demand, work characteristics, traffic control measures, and mobility impacts. A four-set case base schema or domain theory is developed to represent the cases based on the aforementioned characteristics of the problem. It includes a general information set, a problem description set, a solution (or control) description set, and an effects set. To improve the interactivity of the CBR system and its user-friendliness, a hierarchical object-oriented case model is developed for work zone traffic management. The model is implemented into an intelligent decision-support tool to assist traffic agencies in the development of work zone traffic control plans and to better design and manage work zones for increased mobility and safety. Three examples are presented to show the practical utility of the CBR system for work zone traffic management.

CE Database keywords: Highway maintenance; Traffic management; Safety.

DOI: 10.1061/(ASCE)0733-947X(2003)129:2(134)

Introduction

Periodic reconstruction and maintenance of the freeway system is necessary to ensure that it fulfills its long-term purpose of serving the transportation needs of the public efficiently and economically. During the construction and maintenance operations, the normal flow of traffic is disrupted by either a change in the freeway geometry or a temporary freeway closure. Closure of a freeway segment is not a feasible option on most freeways today. Therefore, reconstruction and maintenance operations have to be carried out without entirely closing the freeway segment and in close proximity to traffic flow. Work zones on freeways have to be designed and managed to ensure safety and mobility. The Manual on Uniform Traffic Control Device (FHWA 2000a) provides guidelines for the use of traffic control devices that inform and guide motorists through the work zone with adequate protection for the workers. These guidelines were developed over the years from studies of traffic control devices and their effectiveness in improving work zone safety.

Recently, the Federal Highway Administration (FHWA) reviewed the state of practice in work zone traffic management and found that no uniform and objective procedure exists for quantifying the effects of various factors and determining the life-cycle costs of work zone traffic management plans (FHWA 2000b). They also outlined several steps that should be taken by state and local agencies to satisfy the expectations of the customer (the traveling public). Among the policy, planning, design, and management related steps outlined is the recommendation to "*develop and/or enhance user friendly software to model work zone delay, queues, and crashes; calculate defendable road-user costs and proposed contract time, evaluate proposed changes to the traffic control plan, as well as analyze work zone crashes. All software must be sufficiently flexible to allow for variable parameters to meet unique state/local conditions*" (FHWA 2000b). Towards this end, a Microsoft Excel–based software called QuickZone is being developed for the FHWA for work zone user delay and cost quantification (Mitretek 2000). The software allows planners to model work zones and their associated traffic control plans and provides them with basic delay and queue information that can be used for decision making. A simple macroscopic input-output traffic analysis is adopted in the software to arrive at the estimates. The practical usefulness of the software, however, is limited because it does not maintain a history of previous decisions, nor does it learn from them in reaching a better decision. Furthermore, an input-output analysis assumes that the planner knows the effect of each work zone configuration in terms of the reduction in roadway capacity (maximum number of vehicles that can pass through a roadway segment in one hour under prevailing conditions, expressed as vehicles per hour per lane). This information is usually not available and the planner has to make an educated guess that may or may not be accurate, thus leading to erroneous conclusions.

Case-based reasoning (CBR) is a methodology for storing and retrieving previous design decisions or cases and adapting them to the solution of new problem cases not found in the case base (Leake 1996; Maher and Pu 1997). The CBR approach does not require a low-level physical model of the problem. Rather, in a manner similar to human reasoning and decision making, it uses generic and problem-specific similarity metrics to induce best solutions from previously solved cases. This approach is appropriate for the work zone traffic management problem for the following reasons: (1) accurate mathematical models of work zone traffic flow are not available; (2) there are only a finite number of cases to be considered; and (3) traffic agencies can use previously solved cases to set up the case base and then build it up gradually.

[1]Assistant Professor, Computer Science Dept., Lahore Univ. of Management Sciences, Pakistan.

[2]Professor, Dept. of Civil and Environmental Engineering and Geodetic Science, The Ohio State Univ., 470 Hitchcock Hall, 2070 Neil Ave., Columbus, OH 43210.

Note. Discussion open until August 1, 2003. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on October 23, 2001; approved on March 28, 2002. This paper is part of the ***Journal of Transportation Engineering***, Vol. 129, No. 2, March 1, 2003. ©ASCE, ISSN 0733-947X/2003/2-134–145/$18.00.

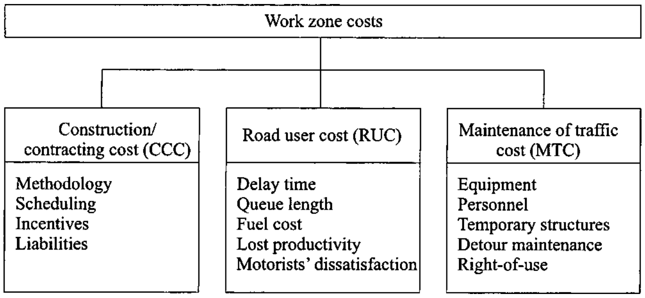


Fig. 1. Freeway construction work zone costs and factors affecting them

Work Zones and Traffic Management

A work zone is a region within an existing freeway's roadway where active maintenance, rehabilitation, and/or reconstruction work is carried out. The freeway is not closed and traffic and freeway work exist in close proximity to each other. A work zone thus represents a spatial and temporal restriction on a freeway's roadway that negatively impacts the normal flow of traffic. The impact appears in the form of increased congestion, travel times, accidents, and a greater level of dissatisfaction among the traveling public. Work zones are designed and managed to minimize these effects and the overall cost.

Work zone costs are often divided into three components: construction/contracting cost (CCC), road user cost (RUC), and maintenance of traffic cost (MTC) (Fig. 1). Construction and contracting cost is the amount charged by the contractor for the work plus any litigation and liability cost. Road user cost is the result of the negative impact of the work zone on the normal flow of traffic. Road user cost can be quantified in several ways, including delay time, queue length, lost productivity, fuel wastage, and pollution. There is also a nonquantifiable aspect to the road user cost, that of dissatisfied travelers. Nonquantifiable parameters, as the name indicate, are those that cannot be readily expressed in numbers. They are categorized under linguistic terms that are understood by traffic engineers. Maintenance of traffic cost is the cost of labor and equipment needed for maintaining traffic through and around work zones. It includes the cost of traffic control devices such as variable message signs, maintenance of alternate routes, construction and maintenance of temporary pavements, and public dissemination of information through mass media advertisements.

Ideally, management of a work zone requires minimization of the total cost. However, from the highway traffic agency's perspective, the road user cost is the most important cost to consider in a work zone project. All other costs are given a lesser priority. The Ohio Department of Transportation (ODOT), for example, has identified four objectives (in no particular preferential order) to rate traffic control plans. These are: (1) to reduce motorist complaints; (2) to maximize corridor capacity; (3) to minimize duration of motorist inconvenience; and (4) to maximize motorist/work safety. These objectives then become the basis for determining the relative effectiveness of new traffic control plans with respect to previously implemented plans for similar work zone conditions. Currently, this procedure is done manually by traffic engineers based on their previous experiences. This research advocates case-based reasoning as an effective approach for formalizing and automating this procedure to achieve greater reliability and efficiency.

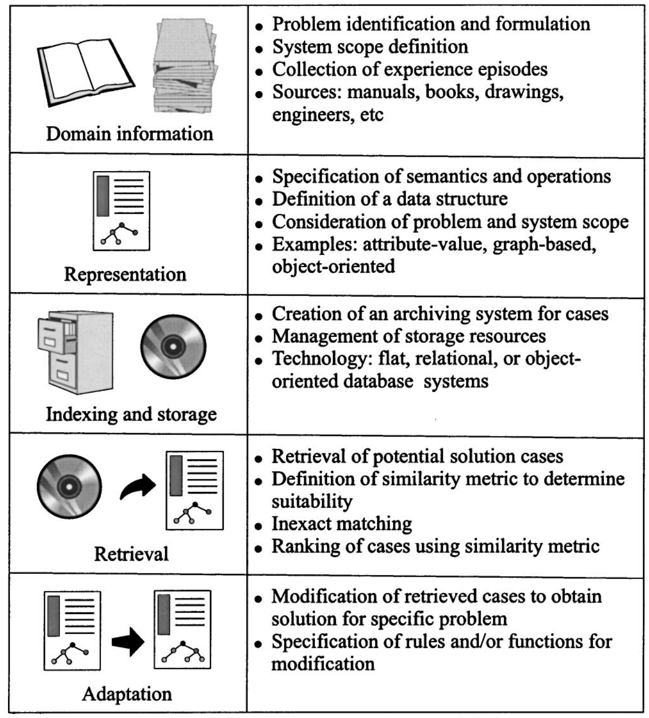


Fig. 2. Elements of case-based reasoning

Case-Based Reasoning

Case-based reasoning (CBR) evolved from cognitive science research into an intelligent problem-solving approach that relies on previous experiences in the form of cases of previously solved similar problems. CBR is a multidisciplinary subject that is viewed with different perspectives in cognitive science, artificial intelligence, and knowledge engineering (Aha 1998). It is loosely based on human reasoning and problem solving, which is essentially experiential and episode based. For example, an experienced traffic engineer can plan a work zone by recalling the knowledge gained from similar scenarios that he or she had solved previously, thus avoiding starting from scratch. CBR can be therefore thought of as a high level model of human reasoning and problem solving, which is the view adopted in cognitive science. In artificial intelligence and knowledge engineering, modeling of human reasoning is not the goal per se but the basis for the development of computational models for the solution of real world problems. Case-based reasoning systems thus mimic human reasoning by retrieving and revising cases from memory (previous experiences) to find solutions for new problems in a given domain.

CBR systems differ from rule-based and model-based systems (Adeli 1998; 1988, 1990a,b; Adeli and Balasubramanyam 1988) in that they require little low-level domain knowledge and rely more on general rules for retrieving and adapting saved solutions. A major drawback of rule-based systems is the difficulty in eliciting knowledge in the form of low-level rules from experts to be used by an inference engine that chains these rules to arrive at a reliable solution. Model-based systems, on the other hand, assume that an accurate mathematical model for the problem exists. This is often not the case for complex engineering problems such as the work zone traffic control problem.

The development of a CBR system requires the specification of several procedures. A schematic description of these proce-

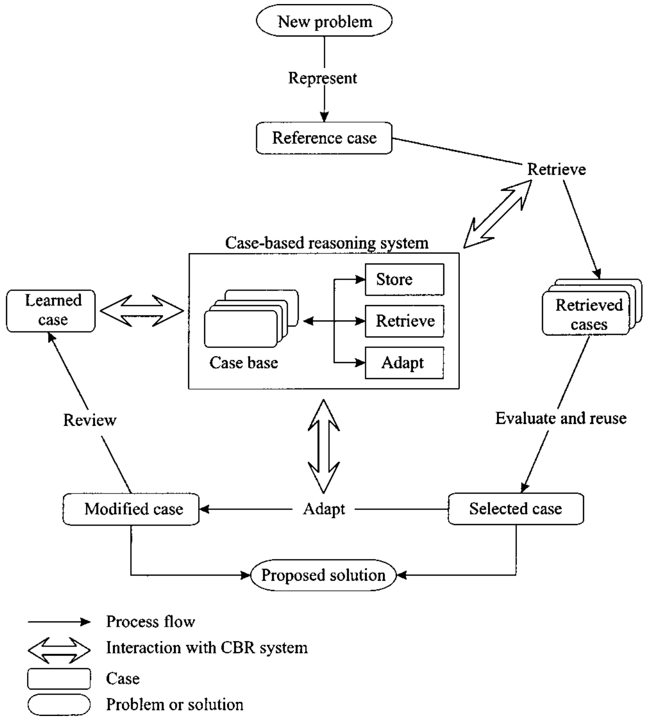


Fig. 3. Typical CBR system processing cycle

dures is given in Fig. 2. These procedures represent typical operations in a CBR system. As such, they may also be thought of as typical CBR system components. Case-based reasoning is a methodology for solving problems and not a specific artificial intelligence technique (Watson 1999). A typical case-based reasoning and problem-solving cycle is shown in Fig. 3. A new problem is first represented into a reference case. This case specifies the problem requirements, which may or may not be complete, and their relative importance. Using this reference case, the CBR system ranks cases in the case base according to their degree of similarity to the reference case. If the retrieved cases do not provide a satisfactory solution, which is usually the case, then they are used as the starting solution to be revised and adapted in order to obtain an improved or satisfactory solution. The retrieval performance of a CBR system improves as the number of reliable cases in the case base increases. Initially, a functional CBR system may have only a few cases in the case base; new cases are consequently added as new problems are solved. This is how *learning* occurs in a CBR system.

CBR systems have been developed for design, planning, decision support, and diagnosis in diverse fields such as engineering, medicine, law, and business (Leake 1996; Maher and Pu 1997; Lenz et al. 1998; Aha 1998). However, the development of a CBR system for work zone traffic planning and management has not been reported in the literature.

Objectives

In this research, a case-based reasoning approach is used for the development of a decision-support system for work zone traffic management with the following objectives:

1. To provide traffic engineers with an intelligent decision-support tool for design of freeway (and divided highway) work zone traffic control plans with the goal of reducing the road user cost (reduced complaints, increased corridor capacity, reduced delays, and improved safety),
2. To develop a case base schema or domain theory to represent the cases based on factors such as work zone layout, traffic demand, work characteristics, traffic control measures, and mobility impacts,
3. To develop work zone traffic control plans that are reliable and defendable,
4. To archive previous experiences of work zone traffic control for quick reference, and
5. To serve as a learning and training tool for work zone traffic control under different work zone scenarios.

Scope and Categorization of Parameters

The scope of applicability of the CBR system for work zone traffic management is defined and parameters involved are categorized in this section.

Work Zone Type

Several types of work zones are used in practice. The scope of the CBR system is limited to temporary stationary work zones on divided highways or freeways. Short-duration or mobile work zones (with durations of less than an hour) are not considered because standard traffic control plans are often adequate for maintaining traffic flow through such work zones. For a given work zone, a separate traffic control plan is developed for each direction of flow independent of flow in the other direction. This simplifies the modeling and understanding of work zone traffic flow by reducing the number of variables to consider.

Work Zone Layout

The CBR system can consider part-width construction (lane merging), lane shifting, and crossover layouts. In part-width construction, one or more lanes are closed to traffic and traffic is merged into the remaining open lanes. Such a layout is usually represented by $[a,b]$ $(a>b)$ where a and b are the number of open lanes before and after the establishment of the work zone, respectively. In lane shifting layout, the number of lanes is not reduced and traffic is shifted around the work zone on temporary pavements or shoulders. No merging operation occurs in a lane shifting work zone layout. Crossover layouts are the combination of lane shifting and lane merging layouts, where traffic is merged and shifted across the median onto lane(s) for travel in the other direction. Thus, the two streams of traffic share the same roadway in close proximity to each other.

Work Characteristics

It has been found that the capacity of a work zone depends on the type and intensity of work (Krammes and Lopez 1994; Dixon et al. 1996). This in turn affects the flow of traffic through work zones. The proximity of heavy equipment, workers, noise, and dust tends to reduce mean speeds through work zones; work of greater intensity produces a greater impact than work of lesser intensity. These factors are considered in the CBR system by qualitative grades of intensity of work specified as part of the description of the work zone scenario.

Traffic Flow Characteristics

Traffic control plans are developed to facilitate the flow of traffic through and around work zones. To develop effective plans, it is necessary to have the highway segment's traffic flow characteristics, such as flow rate, traffic composition, and driver behavior. The traffic demand that needs to be handled can be specified by the hourly flow rate on the highway segment prior to the establishment of the work zone. The percentage of trucks gives an indication of the traffic stream's composition, which in turn gives an indication of flow characteristics such as average speed. The familiarity of the drivers with the highway corridor also has a significant impact. This can be captured in a qualitative manner by categorizing highways as urban, suburban, or rural. The CBR system can consider all these factors for analysis of work zone traffic flow. The hourly flow rate is required, while the others are optional if reliable data is available.

Phases of Work

A work zone may go through several phases over its lifetime. Work enters a new phase whenever any of the parameters defining the work zone scenario changes. Changes in work zone scenarios are analyzed by creating a new problem description and developing traffic control plans for each one separately. The CBR system considers the duration of a phase to determine the time-dependent impact of the work zone scenario.

Traffic Control Measures

It is assumed that the requirements of the Manual on Uniform Traffic Control Devices (FHWA 2000a) are followed for all traffic control plans. To improve mobility further, the traffic agency can take further measures such as providing signed alternate routes, advanced roadside warning and informative messages, and updates on traffic conditions through the mass media, and by posting reduced speed limits in the work zone. These factors are considered in the CBR system in a qualitative manner. Note that the impact of these measures will depend on traffic flow characteristics in the given highway, such as flow rate and driver behavior.

Road User Cost

Road user cost is the determining criterion for the selection of a traffic control plan for a work zone. Quantifying actual cost incurred by road users is difficult. Therefore, indirect measures of the negative impacts of work zones are usually used. As a measure of motorist inconvenience, the CBR system uses the quantitative measures of maximum queue length and delay time that motorists can experience as a result of a given work zone traffic control plan. Furthermore, the CBR system considers motorists' complaints, corridor capacity, and safety in a qualitative manner. These criteria correspond to the four objectives identified by ODOT for the design of work zone traffic control plans. The CBR system works even when only one of these values is given for a work zone scenario.

Four-Set Case Model for the Work Zone Traffic Management Domain

A case model or domain theory is a template for collection of information that captures a problem-solution episode. In general, this information is usually partitioned into two sets: a problem set and a solution set. The problem set contains information that describes the problem whose solution is desired. This information

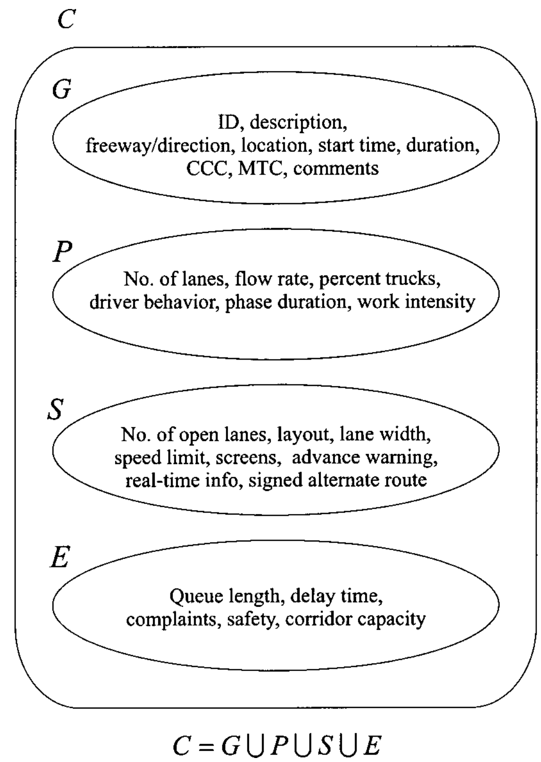


Fig. 4. Four-set case model for CBR system for work zone traffic management

uniquely identifies the case in the case base. The solution set contains information that describes the solution chosen for the problem.

Considering the scope of the CBR system for work zone traffic management, a two-set case model is neither adequate nor appropriate. Each case must contain all the information needed for case-based reasoning plus the information required for maintaining complete records of previous experiences for administrative purposes. Furthermore, the outputs of the system must include information on the effects of the traffic control plan chosen for a given problem description. For these reasons, in this research we create a four-set case model for work zone traffic management consisting of a general information set (G), a problem description set (P), a solution (or control) description set (S), and an effects set (E). Mathematically, a case is defined as the union of the four nonoverlapping or disjoint sets as follows:

$$C = G \cup P \cup S \cup E \tag{1}$$

where $\cup$ = the set union operator (Fig. 4).

The general set contains information that identifies and describes the experience episode for future reference. Any useful information beyond that needed for the operation of the CBR system is included in this set so that a complete record of the previous experience episode is maintained in the case. The problem set contains information that defines the constants of the work zone traffic control problem. This information is known to the traffic engineer from construction plans and traffic studies and represents work zone conditions. Information in this set includes number of lanes, flow rate, duration of work, and intensity of work.

The solution or control set contains information on the work zone layout and traffic control measures adopted for the mitigation of traffic congestion. This information defines the solution, or

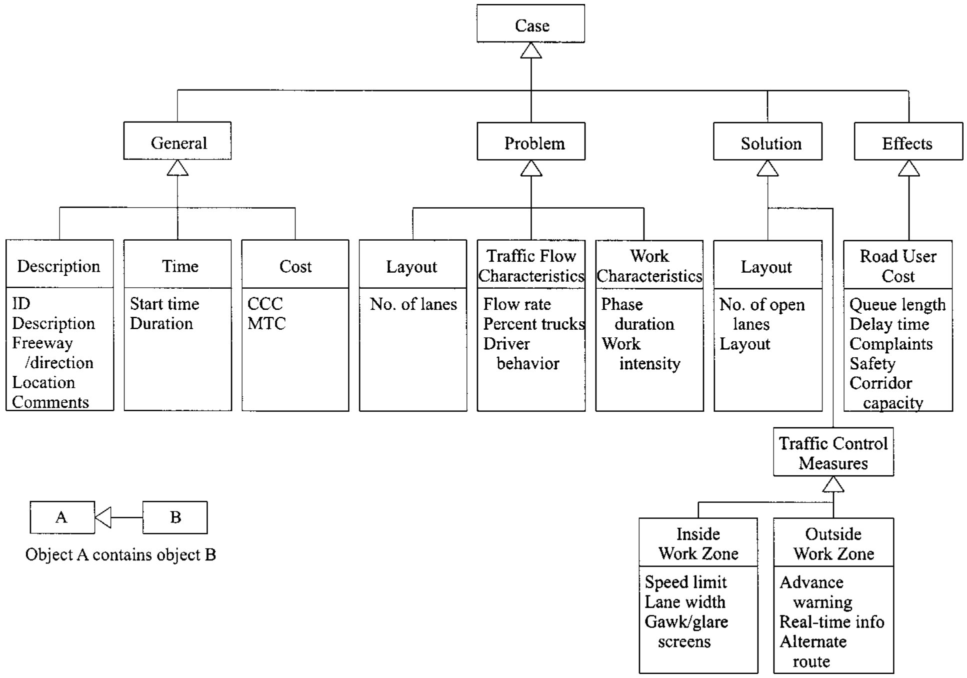


Fig. 5. Object-oriented case model for CBR system for work zone traffic management

the traffic control plan, for the work zone defined in the problem set. Information in the solution set includes number of open lanes, work zone layout, and traffic control measures such as advance motorists' warning and signed alternate routes. The effects set contains information about the impacts on the traffic in the work zone. This information forms the criteria for the selection of one case over another as a solution for a given work zone traffic control problem.

In the case model for the work zone traffic management, each case is uniquely identified by the union of the problem (P) and solution (S) sets. Thus, two cases in the case base can have identical problem sets; however, their solution sets must differ. This situation may represent two experience episodes where the work zone traffic control problem is identical but a different traffic control plan is adopted for each with possibly different impacts. When querying the system, the traffic engineer can specify as much of the information in the problem and solution sets as known. The more information the traffic engineer provides, the more specific will be the cases retrieved by the CBR system. It should be pointed out that it is not necessary to specify all the information in the problem set, because the CBR approach does not require exact matching for retrieval.

Eq. (1) defines a case as a set of information. The case base can then be defined as the union of all the cases $C_i = G_i \cup P_i \cup S_i \cup E_i$

$$Z = \cup_i C_i \tag{2}$$

such that

$$C_i \neq C_j \Leftrightarrow P_i \cup S_i \neq P_j \cup S_j' \quad \forall i,j,i \neq j \tag{3}$$

Eq. (3) ensures that no two cases in the case base have the same problem and solution sets and all cases are unique. The case base given by the set Z captures the domain knowledge needed for solving the problem. The effectiveness of the CBR system increases as the number and diversity of cases in the case base increases, encompassing the entire knowledge domain defined by its scope of applicability. The CBR system, however, can work even with only a few cases in the case base.

Table 1. Attribute-Value Representation of Information in *General* Object

Name	Description	Type	Value representation	Example
ID	Case identification	Choice	Free-form alphanumeric	OH-5235
Description	Brief description of work zone traffic control project	Choice	Free-form alphanumeric	Resurfacing of southbound lane
Freeway/direction	Freeway identification number and direction	Choice	Designation/[NB, SB, EB, WB][a]	I-71/NB
Location	Geographical location of freeway	Choice	County, city	Franklin, Columbus
Start time	Start time of project	Choice	Year, month	2000, 02
Duration	Duration of project	Number	Days	30
CCC	Construction/contracting cost	Number	Thousand dollars	25,000
MTC	Maintenance of traffic cost	Number	Thousand dollars	500
Comments	Additional comments	Choice	Free-form alphanumeric	Completed successfully

[a]NB=northbound; SB=southbound; EB=eastbound; WB=westbound.

Table 2. Attribute-Value Representation of Information in the *Problem* Object

Name	Description	Type	Value representation	Example
Number of lanes	Number of open lanes prior to creation of work zone	Number	Positive integer	3
Flow rate	Average flow at work zone site	Number	Vehicles/hour/lane	1500
Percent trucks	Percentage of heavy vehicles or trucks in traffic stream	Number	Percent	5
Driver behavior	Classification of driver behavior	Choice	[Urban, rural]	Urban
Phase duration	Duration for work phase	Number	Hours	4
Work intensity	Classification of work intensity	Choice	[High, moderate, low]	Moderate

Hierarchical Object-Oriented Case Model

The representation of a case as a union of information sets is most appropriate for the design of a CBR system. This representation partitions the variables involved in the problem according to their use in the CBR system: input, output, indexing, retrieval, and adaptation. However, this partitioning is not appropriate for human comprehension and the user friendliness of the CBR system. Over the years, traffic engineers have developed a body of knowledge for work zone traffic control that categorizes information in a manner similar to that presented in a previous section. This categorization is based on key elements or components of the work zone traffic control problem and is generally more specialized than the four-set categorization defined for the set representation of the case model. A case model that provides such a level of detail is useful for the design of an effective user interface for the CBR system. An object-oriented representation is used to create such a user interface.

A hierarchical object-oriented case model is developed for the CBR system for work zone traffic management (Fig. 5). A case in the system, represented by a ''*Case*'' object, uses four lower level objects, ''*General*,'' ''*Problem*,'' ''*Solution*,'' and ''*Effects*,'' corresponding to the four sets defined in the set model of the case. The *General* object uses three lower level objects, *Description*, *Time*, and *Cost*, which collectively encapsulate general information needed to keep a complete record of the experience episode. The *General* object can own additional objects depending on the information needs of the user.

The *Problem* object uses three lower level objects, *Layout*, *Traffic Flow Characteristics*, and *Work Characteristics*. These objects encapsulate the work zone traffic control problem or the pre-existing geometry and flow conditions for which a traffic control plan is desired. The *Solution* object encapsulates the traffic control plan. It uses two lower level objects: *Layout* and *Traffic Control Measures*. The *Layout* object encapsulates information about the geometric conditions after the establishment of the work zone, while the *Traffic Control Measures* object encapsulates the steps taken to alleviate traffic congestion. Work zone traffic control measures are often divided into those taken inside the work zone and those taken outside the work zone. The lowest objects, *Inside Work Zone* and *Outside Work Zone*, capture this division of information. Traffic control measures taken inside a work zone include imposing speed limits, widening lanes, and erecting gawk screens, while those taken outside the work zone include warning motorists in advance and diverting traffic through alternate routes. The *Effects* object encapsulates information on the effects of the traffic control plan, which is essentially the road user cost. The *Road User Cost* object describes the impact of the traffic control plan on motorists.

The most specialized objects in the object-oriented case model for work zone traffic management (the leaf nodes in Fig. 5) define the categories readily understood by traffic engineers. Information in these categories is merged to form the four-set case model used by the CBR system.

Case Representation

In the case models presented in the previous section, a case is defined as a collection of information objects. The information in the objects is identified by linguistic terms that are generally understood by humans but are imprecise for information processing. Information or knowledge representation involves the specification of semantics to information entities, which enables machines to use well-defined operations to process them.

Since cases and objects in the CBR system for work zone traffic management are a collection of facts rather than rules or functions, an attribute-value scheme is used for information representation. An attribute-value representation of information is defined by three elements:

- An attribute or field name that identifies the information entity and gives it a meaning that can be understood by humans,
- A type that specifies the type of the attribute, and
- A value taken from the domain that specifies the current instantiation of the attribute.

Common attribute types include choice (free-form text), alphabetic, number, integer, and positive number. A range can also be

Table 3. Attribute-Value Representation of Information in *Solution* Object

Name	Description	Type	Value representation	Example
Number of open lanes	Number of open lanes after creation of work zone	Number	Positive integer	2
Layout	Work zone layout or configuration	Choice	[Merge, shift, crossover]	Merge
Speed limit	Posted speed limit within work zone	Number	1.61×km/h (mi/h)	45
Lane width	Width of lanes within work zone	Number	0.305×m (ft)	11
Screens	Gawk/glare screens to prevent driver distractions	Choice	[Yes, no]	No
Advance warning	Advance warning of work zone before exits and alternate routes	Choice	[Yes, no]	Yes
Real-time info	Real-time info on traffic congestion ahead of work zone	Choice	[Yes, no]	No
Signed alternate route	Signed alternate routes ahead of work zone	Choice	[Yes, no]	Yes

Table 4. Attribute-Value Representation of Information in *Effects* Object

Name	Description	Type	Value representation	Example
Queue length	Maximum queue length observed during work phase	Number	1.61×km (mi)	2
Delay time	Maximum delay time experience during work phase	Number	Vehicle-hours	2,500
Complaints	Amount of motorists' complaints	Choice	[High, medium, low]	Low
Safety	Level of motorist and worker safety	Choice	[High, medium, low]	High
Corridor capacity	Reduction in corridor capacity	Choice	[High, medium, low]	Medium

specified to further constrain and elucidate the domain defined by the type. A range specification may be a list of values, a range of values, a hierarchy of values, or values of a certain unit.

The attribute-value representation A of an information entity can be written as a 3-tuple variable:

$$A=\{name,type,value\} \tag{4}$$

Given an attribute-value representation A the elements are defined by the functions Name(A) = *name*, Type(A) = *type*, and Value(A) = *value* = v. Therefore, a case C_i in the CBR system for work zone traffic management can be represented by a collection of attribute-value representations of all the information entities it contains. This can be written as

$$C_i=\{A_1^i,A_2^i,A_3^i,\ldots,A_N^i\} \tag{5}$$

where A_j^i=jth attribute-value representation in case i and N =total number of attributes in a case. The *name* and *type* elements of a given attribute-value representation i ($i=1,N$) are identical in all cases in the case base; the *value* elements, however, may be different. The attribute-value representations of the information entities that constitute a case in the CBR system for work zone traffic management, corresponding to the *General*, *Problem*, *Solution*, and *Effects* sets, are defined in Tables 1–4. Only two types of values are used for representation: choice and number.

Similarity Measures

The degree of similarity between numeric attribute i of two cases j and k is defined as

$$\text{Similarity}(A_i^j,A_i^k)=\frac{\min(|v_i^j|,|v_i^k|)}{\max(|v_i^j|,|v_i^k|)} \tag{6}$$

where $v_i^j=\text{value}(A_i^j)\neq 0$ and $|\cdot|$ denotes the absolute value. In the CBR system for work zone traffic management, all values of numeric attributes are nonzero and positive. Thus, Eq. (6) computes the degree of similarity as the ratio of the minimum value to the maximum value, which ranges from greater than 0 to 1.

The degree of similarity between choice (free-form text) attribute type i of two cases, j and k, is defined by the following rule:

$$\text{IF}\quad (v_i^j \ \text{appears in}\ v_i^k)\ \text{OR}\ (v_i^k\ \text{appears in}\ v_i^j)$$
$$\text{THEN}\ \ \text{Similarity}(A_i^j,A_i^k)=1\quad \text{ELSE}\ \ \text{Similarity}(A_i^j,A_i^k)=0 \tag{7}$$

Since the choice type represents free-form text, it may consist of numbers, alphabets, and special characters (such as spaces). Note that the similarity operations are commutative; that is, $\text{Similarity}(A_i^j,A_i^k)=\text{Similarity}(A_i^k,A_i^j)$.

Case Retrieval

An interaction with a CBR system starts with the formulation of a query that describes a situation for which a solution is desired. Based on this query, the system retrieves cases from the case base as potential solutions to the problem. The retrieval process is guided by the degree of similarity (or match) of the query to the cases in the case base. In the CBR system for work zone traffic management, the query consists of two components: a reference case and a weight vector. The reference case R is defined as

$$R=\{A_1,A_2,A_3,\ldots,A_N\} \tag{8}$$

This equation is similar to Eq. (5). Thus, a reference case has the same collection of attribute-value representations as other cases in the CBR system. The traffic engineer using the CBR system inputs values for the attributes in the reference case to describe the work zone scenario. The weight vector w_i ($i=1,N$) attaches an importance to the similarity of each attribute in the retrieval process. The suitability of the cases in the case base as solutions to the query is determined by a case or global similarity measure. This is computed as the weighed sum of the similarities of the respective case and reference case values. The case similarity function for case i as compared with a given reference case R is defined as

$$\text{Similarity}(C_i,R)=\frac{\sum_{j=1}^{N} w_j\times\text{Similarity}(A_j^i,A_j^R)}{\sum_{j=1}^{N}|w_j|} \tag{9}$$

Case similarity scores range from 0 to 1, where 0 indicates no similarity while 1 denotes full similarity. Based on the case simi-

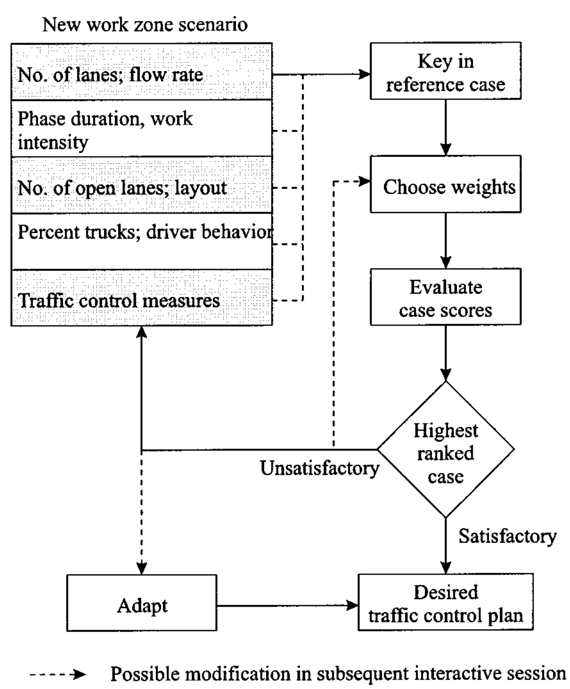


Fig. 6. Procedure for creation of work zone traffic control plans using CBR system for work zone traffic management

Microsoft Excel - workzone.xls

	A	B	C	D	E	F	G	H	I	J	K
1	**Induce-It**	Copyright ©1992-1999 Inductive Solutions, Inc.									
2											
3											
4		Types:	c	c	c	c	c	n	n	n	c
5		Maps:									
6		Field Names:	ID	Description	Freeway	Location	Start time	Duration	CCC	MTC	Comments
7		**Weights:**	1	1	1	1	1	1	1	1	1
8		**Reference:**									
9											
10	**Scores**	**Cases**									
11	0.992	Case1	OH-0142	Resurfacing	I-71/NB	Franklin	1998/02	25	2500	25	
12	0.986	Case2	OH-0258	Pavement repair	I-75/NB	Morrow	1999/06	5	500	15	
13	0.986	Case3	OH-0268	Bridge work	I-71/NB	Delaware	1999/08	20	250	15	
14	0.984	Case4	OH-0555	Resurfacing	I-71/SB	Delaware	2000/07	35	2000	20	
15	0.987	Case5	OH-0233	Pavement rehabilitation	I-70/EB	Franklin	1999/04	60	4500	30	
16	0.982	Case6	OH-0325	Resurfacing	I-70/EB	Franklin	2000/03	15	1500	20	
17	0.992	Case7	OH-0382	Pavement marking	I-71/NB	Franklin	2000/04	5	250	5	
18	0.980	Case8	OH-0422	Lane addition	I-70/WB		2000/06	75	5500	25	
19	0.984	Case9	OH-0501	Utility work	I-71/SB		2000/06	5	35	5	
20	0.984	Case10	OH-0155	Pavement joint work	I-71/SB	Delaware	1998/03	15	25	10	
21	0.988	Case11	OH-0298	Bridge work	I-80/EB		1998/12	15	250	5	
22	1.000	Case12	OH-0333	Pavement marking	I-80/EB		1999/04	10	10	3	
23	0.986	Case13	OH-0482	Resurfacing	I-77/NB		2000/05	35	125	10	
24	0.984	Case14	OH-0186	Stripping	I-77/SB		1998/04	10	55	5	
25	0.986	Case15	OH-0208	Pavement rehabilitation	I-71/NB		1999/03	45	325	25	
26	0.980	Case16	OH-0342	Pavement repair	I-71/SB		2000/03	25	200	15	
27	1.000	Case17	OH-0329	Culvert work	I-77/SB		2000/03	10	80	5	
28	0.998	Case18	OH-462	Pavement marking	I-77/SB		2000/05	5	15	3	
29	0.985	Case19	OH-444	Utility work	I-80/WB	Cuyahoga	2000/04	5	25	5	
30	0.977	Case20	OH-0218	Lane addition	I-70/EB	Franklin	1999/08	55	480	20	
31											

Fig. 7. CBR system user interface showing reference case, weights, case scores, and sample case base for work zone traffic management corresponding to *General* object

larity scores, the cases in the case base are ranked and presented to the user. Cases with the largest score represent potential solutions for the problem at hand.

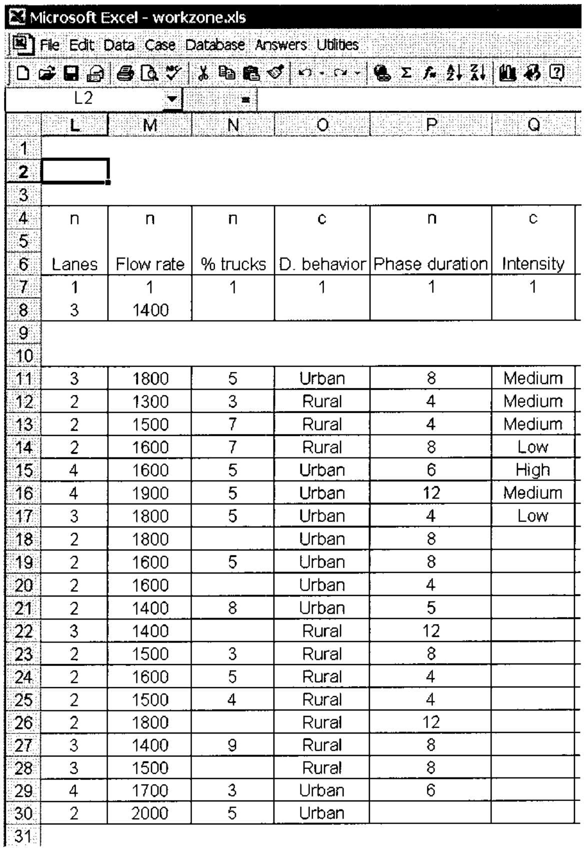
Creation of the Case Base

The CBR system for freeway work zone management has been implemented in Induce-It, a software shell for developing case-based reasoning systems (Inductive Solutions 2000). Induce-It is based on the Microsoft Excel spreadsheet software system and relies on its user interface, database, and programming capabilities to provide an environment for developing and using a CBR system. Induce-It provides built-in capabilities for case representation, indexing, storage, retrieval, and adaptation, allowing the developer to concentrate on domain information collection and problem formulation. Cases are represented as a sequence of attribute-value pairs. Induce-It supports several numeric and textual field types including number, choice (free-form text), and user-specified. A specific region in the spreadsheet is reserved for the case base, where cases appear in rows while case field values appear in columns.

Based on the case models presented in the preceding sections, a prototype CBR system for work zone traffic management is developed using Induce-It. The case base of the CBR system presently includes twenty cases representing common work zone scenarios and their corresponding traffic control plans. The cases were created from information obtained from the Ohio Department of Transportation. The information consisted primarily of qualitative data such as work zone classification, traffic control measures, planning goals, and development procedures. The quantitative data used in the cases such as the freeway traffic flow rate (in the absence of the work zone), maximum queue length, and maximum delay time are derived from human experience of work zone traffic control. The sample case base is sufficient for testing the prototype system and can be extended easily as new cases become available.

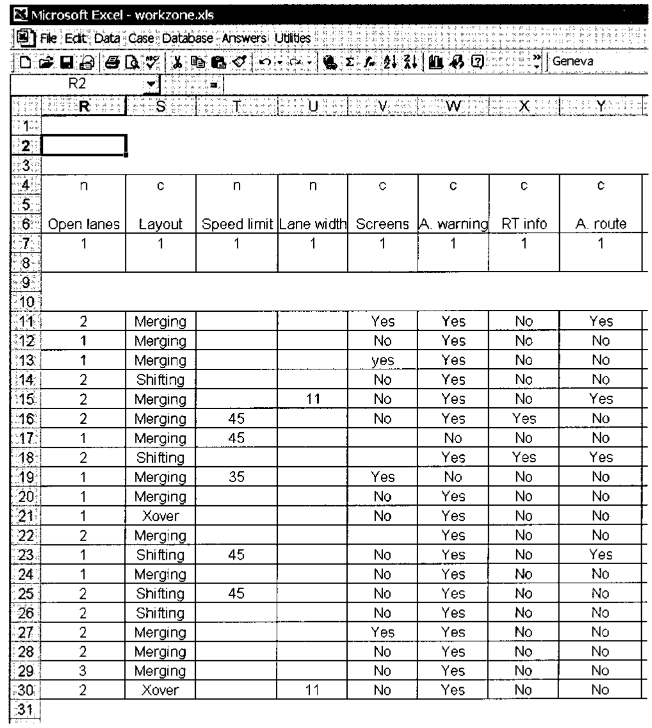
Creation of Work Zone Traffic Control Plans Using the CBR System

The flow chart of steps involved for creation of a suitable work zone traffic control plan using the proposed CBR system is shown in Fig. 6. When a traffic engineer wants to create a traffic control plan for a given work zone scenario, he starts with some basic fixed information about the work zone under consideration, such as the number of lanes and flow rate. This information is fed into the CBR system by responding to queries made by the system. This is done in an iterative manner through a number of interac-



	L	M	N	O	P	Q
4	n	n	n	c	n	c
6	Lanes	Flow rate	% trucks	D. behavior	Phase duration	Intensity
7	1	1	1	1	1	1
8	3	1400				
11	3	1800	5	Urban	8	Medium
12	2	1300	3	Rural	4	Medium
13	2	1500	7	Rural	4	Medium
14	2	1600	7	Rural	8	Low
15	4	1600	5	Urban	6	High
16	4	1900	5	Urban	12	Medium
17	3	1800	5	Urban	4	Low
18	2	1800		Urban	8	
19	2	1600	5	Urban	8	
20	2	1600		Urban	4	
21	2	1400	8	Urban	5	
22	3	1400		Rural	12	
23	2	1500	3	Rural	8	
24	2	1600	5	Rural	4	
25	2	1500	4	Rural	4	
26	2	1800		Rural	12	
27	3	1400	9	Rural	8	
28	3	1500		Rural	8	
29	4	1700	3	Urban	6	
30	2	2000	5	Urban		

Fig. 8. Sample case base for work zone traffic management corresponding to *Problem* object



	R	S	T	U	V	W	X	Y
4	n	c	n	n	c	c	c	c
6	Open lanes	Layout	Speed limit	Lane width	Screens	A. warning	RT info	A. route
7	1	1	1	1	1	1	1	1
11	2	Merging			Yes	Yes	No	Yes
12	1	Merging			No	Yes	No	No
13	1	Merging			yes	Yes	No	No
14	2	Shifting			No	Yes	No	No
15	2	Merging		11	No	Yes	No	Yes
16	2	Merging	45		No	Yes	Yes	No
17	1	Merging	45			No	No	No
18	2	Shifting				Yes	Yes	Yes
19	1	Merging	35		Yes	No	No	No
20	1	Merging			No	Yes	No	No
21	1	Xover			No	Yes	No	No
22	2	Merging				Yes	No	No
23	1	Shifting	45		No	Yes	No	Yes
24	1	Merging			No	Yes	No	No
25	2	Shifting	45		No	Yes	No	No
26	2	Shifting			No	Yes	No	No
27	2	Merging			Yes	Yes	No	No
28	2	Merging			No	Yes	No	No
29	3	Merging			No	Yes	No	No
30	2	Xover		11	No	Yes	No	No

Fig. 9. Sample case base for work zone traffic management corresponding to *Solution* object

tive sessions until a satisfactory solution case is obtained or a retrieved case is adapted to obtain a desired solution.

Initially the reference case is created with the minimum information needed to describe the work zone situation, that is, the number of lanes and the flow rate. This ensures that a wide spectrum of cases is retrieved by the system. If after evaluating the retrieved cases based on the case scores no suitable solution is found, the reference case is modified in the subsequent interactive sessions by adding more information known about the work zone scenario. In general, the reference case is modified in the sequence shown in the top left corner of Fig. 6, where at each subsequent interactive session the information in the next lower box is added to the reference case. This procedure ensures that the solution is narrowed down gradually and minimizes the possibility of missing good solutions by first starting with minimum required input.

The traffic engineer using the CBR system can use his judgment to assign weights to various attributes. The value of each weight indicates the significance of the corresponding attribute. For example, if it is desired that at least two lanes be open, then the number of open lanes attribute should be given a larger weight. Also, a weight can indicate the reliability of a given value. For example, if the flow rate is not known accurately then a lower weight should be assigned to it. In general, the weights need not be changed from one interactive session to the next. However, the CBR system user can modify them for the same reference case to tune the output of the system.

The retrieved cases are compared according to their case similarity scores computed by the CBR system. A higher score indicates a closer match to the reference case and the weights inputted by the user. In addition to this automatic suitability measure, the CBR system user can also evaluate the retrieved cases for their impacts on motorists, the number and type of traffic control measures, and the maintenance of traffic cost. This evaluation will guide the traffic engineer to modify the reference case and the associated weights, accepting a case as the desired solution or modifying a case to obtain an improved solution.

Case adaptation is attempted after several interactive sessions yield no desired solution from the case base. Using the retrieved cases as a guide, the traffic engineer can modify them to arrive at a desirable solution. This solution may then be included in the case base for future perusal.

Illustrative Examples

In this section, the CBR system for work zone traffic management is used to solve three examples. Figs. 7–10, considered side by side, show the CBR system's user interface. They display the attribute-value representation of the information, the reference case, the weights, and the case similarity scores. Figs. 7–10, respectively, show the portion of the case base corresponding to the *General*, *Problem*, *Solution*, and *Effects* objects of the case model. Each case is displayed in a separate row, starting from row 11. The field names and values appear in columns, starting from column C. The reference case is defined in row 8 and the weights indicating the relative importance of the values in the reference are specified in row 7. The suitability of the cases in the case base as potential solutions to the reference case is indicated by the case score, displayed in column A (Fig. 7).

Example 1

This example illustrates the use of the CBR system as a decision-support tool for creation of a work zone traffic control plan.

Microsoft Excel - workzone.xls

File Edit Data Case Database Answers Utilities

Z2

	Z	AA	AB	AC	AD
1					
2					
3					
4	n	n	c	c	c
5					
6	Queue length	Delay time	Complaints	Safety	C. capacity
7	1	1	1	1	1
8					
9					
10					
11	1	750	Low	High	High
12	2	1000		High	
13	0.5				
14	0.5	500			
15	0.5	300	High		
16	2	400			
17	1		Medium	High	High
18	0.5				
19	1		Low	High	
20	1.5	700			
21	3				
22	0.5	200		Medium	
23	4		High		
24	1			High	
25	2	500	Medium		
26	1		Medium		
27	2			High	Medium
28	0.5				
29	1			High	High
30	1		Low	High	
31					

Fig. 10. Sample case base for work zone traffic management corresponding to *Effects* object

Given the description of the work zone scenario as defined by a reference case, the traffic engineer uses the CBR system in the manner shown in Fig. 6 to retrieve the most relevant case(s) from the case base. The work zone scenario (reference case) is described in Table 5. The freeway has three lanes, each carrying an average flow of 1,400 vehicles per hour. Each phase of construction lasts for six hours and it is of medium intensity. These are the constants of the work zone scenario for which a traffic control plan is to be developed. In addition to these constants, it is also desired that two lanes be kept open at all times, that the layout be of merging type, and that a signed alternate route be provided to avoid excessive congestion. This work zone scenario is typical for lane resurfacing projects.

The CBR system is consulted in three interactive sessions. The reference case attribute values and weights chosen for each interactive session and their corresponding case similarity scores are summarized in Tables 5 and 6, respectively. In the first interactive session, the reference case is created with the values for the number of lanes and flow rate only, and each is given equal importance. As seen from Table 6, two cases, Cases 12 and 17, match exactly with the reference case, with a similarity score of 1. This scenario, however, is too general and many work zone scenarios have these characteristics but may require different traffic control plans because of differences in other characteristics.

In the second interactive session, the values for the work phase duration and work intensity are added to the reference case. The weights are modified to reflect the greater relative importance of flow rate and number of lanes in the choice of a traffic control plan. The phase duration is given more importance than the work intensity because the former has a more significant impact on the work zone traffic as compared with the latter. In general, the longer the duration of the work zone, the greater the extent of the congestion. This congestion, however, does not increase without bound, as motorists tend to change their driving habits and reduce demand at the work zone site. For this second interactive session, Case 17 has the highest score followed closely by Case 18. As seen from Figs. 8 and 9, these two cases have similar work zone scenarios and traffic control solutions even though they are for different types of construction work (Case 17 is for culvert work and Case 18 is for pavement marking). However, the minor differences that exist in the problem and solution descriptions of these two cases result in a significant difference in the impacts on traffic. One has a queue length of 3.22 km (2 mi) and the other has a queue length of 0.81 km (0.5 mi) (column Z in Fig. 10). For this reason the third interactive session is made more specific by adding the values for number of open lanes, layout, and alternate route to the reference case (Table 5). These values represent the desired characteristics of the traffic control plan that the traffic engineer feels can reduce traffic impacts. Case 1 (presented in row 11 of Figs. 7–10) has the highest score in this interactive session (Table 6) and thus provides the best traffic control plan for the given work zone scenario.

Example 2

The CBR system for work zone traffic management can also be used for information retrieval and engineer training. For this purpose, a reference case is created that contains values desirable in the retrieved cases. The weights are normally all set equal to 1. Suppose the engineer wants to study all work zone scenarios that have a merging layout from four lanes to two lanes. To retrieve all such cases, a reference case is created with number of lanes set to 4, number of open lanes set to 2, and layout set to "Merging."

Table 5. Reference Case (Work Zone Scenario) and Weights for Example 1

		Weights		
Attribute name	Value	Interactive session 1	Interactive session 2	Interactive session 3
Number of lanes	3	1	2	2
Flow rate	1,400 vph/lane	1	2	2
Phase duration	6	NS	1.5	1.5
Work intensity	Medium	NS	1	1
Number of open lanes	2	NS	NS	2
Layout	Merging	NS	NS	1.5
Signed alternate route	Yes	NS	NS	1

Note: NS = no value is specified; vph = vehicles per hour.

Table 6. Case Scores for Illustrative Examples

Case	Example 1: Interactive session 1	Example 1: Interactive session 2	Example 1: Interactive session 3	Example 2	Example 3
Case 1	0.992	0.973	0.974	0.991	0.964
Case 2	0.986	0.957	0.897	0.964	0.929
Case 3	0.986	0.957	0.897	0.964	0.929
Case 4	0.984	0.925	0.850	0.982	0.964
Case 5	0.987	0.943	0.945	1.000	0.929
Case 6	0.982	0.942	0.913	1.000	0.964
Case 7	0.992	0.936	0.877	0.973	0.929
Case 8	0.980	0.951	0.907	0.982	0.929
Case 9	0.984	0.958	0.897	0.964	0.929
Case 10	0.984	0.954	0.893	0.964	0.929
Case 11	0.988	0.970	0.862	0.964	0.964
Case 12	1.000	0.975	0.945	0.991	0.929
Case 13	0.986	0.961	0.885	0.964	0.964
Case 14	0.984	0.954	0.893	0.964	0.929
Case 15	0.986	0.957	0.881	0.982	0.929
Case 16	0.980	0.939	0.864	0.982	0.929
Case 17	1.000	0.988	0.957	0.991	0.929
Case 18	0.998	0.983	0.953	0.991	0.929
Case 19	0.985	0.972	0.921	0.988	0.929
Case 20	0.977	0.958	0.882	0.982	0.964

The case similarity scores for this example are given in Table 6. Cases 5 and 6, with a case score of 1 match the reference case. Note that for such training information retrieval, only cases with scores of 1 are considered, because exact matches are desired.

Example 3

An advantage of CBR systems for knowledge engineering is that the case base can be developed incrementally and easily by the end user. A fully functional CBR system may have only a few cases initially; the user can add more as he or she encounters new problems not found in the case base. To illustrate this, suppose the user wants to develop a traffic control plan for a 3-to-1 crossover layout in a rural location. Interacting with the system with reference case values of 3 for number of lanes, 1 for number of open lanes, ''Xover'' for layout, and ''Rural'' for driver behavior produces the case similarity scores shown in the last column of Table 6. No exact matches are found. Also, the case similarity scores have a narrow spread with no single case dominating the others. This indicates that a satisfactory solution case does not exist in the case base. For situations like these, the user can develop a traffic control plan from scratch (aided by the cases in the case base) and then add the new case to the case base for future perusal.

Concluding Remarks

Traffic agencies are faced with the challenge of planning, designing, and operating work zones that maximize safety and minimize motorists' inconvenience. The most pressing need is to alleviate excessive congestion by developing work zone traffic control plans that efficiently handle traffic flow through and around work zones. Presently, no rigorous procedures exist for the development of work zone traffic control plans. In this research, a case-based reasoning system is developed as an intelligent decision-support tool to assist traffic engineers in the development of work zone traffic control plans. The CBR system developed in this research is the first decision support tool to help traffic engineers create work zone traffic control plans.

The effectiveness of a work zone traffic control plan is measured by the delay experienced by motorists and/or the length of queue formed on the upstream side. To improve objectivity and reliability of traffic control plans, a multiparadigm computational model is currently being developed that maps traffic flow and work zone characteristics to delay time and queue length. The model will be integrated into the CBR system presented in this article.

Acknowledgment

This paper is based on a research project sponsored by the Ohio Department of Transportation (ODOT) and Federal Highway Administration. The assistance of Ken Linger and Max Braxton in providing ODOT documentation is greatly appreciated.

References

Adeli, H., ed. (1988). *Expert systems in construction and structural engineering*, Chapman and Hall, New York.

Adeli, H., ed. (1990a). *Knowledge engineering, volume one: fundamentals*, McGraw-Hill, New York.

Adeli, H., ed. (1990b). *Knowledge engineering, volume two: applications*, McGraw-Hill, New York.

Adeli, H., and Balasubramanyam, K. V. (1988). *Expert systems for structural design—a new generation*, Prentice-Hall, Englewood Cliffs, N.J.

Aha, D. W. (1998). ''Omnipresence of case-based reasoning in science and application.'' *Knowledge-based systems*, 11, 261–273.

Dixon, K. K., Hummer, J. E., and Lorscheider, A. R. (1996). ''Capacity for North Carolina freeway work zones.'' *Transportation Research Record*, 1529, Transportation Research Board, Washington, D.C. 27–34.

Federal Highway Administration (FHWA) (2000a). *Manual on uniform traffic control devices, millennium edition of MUTCD*, ⟨http://mutcd.fhwa.dot.gov/⟩.

Federal Highway Administration FHWA (2000b). *Meeting the customer's needs for mobility and safety during construction and maintenance operations*, ⟨http://www.fhwa.dot.gov/reports/bestprac.pdf⟩.

Inductive Solutions. (2000). *Induce-It user manual*, ⟨http://www.inductive.com/⟩.

Krammes, R. A., and Lopez, G. O. (1994). ''Updated capacity values for short-term freeway work zone lane closures.'' *Transportation Research Record 1442*, Transportation Research Board, Washington, D.C., 49–56.

Leake, D., ed. (1996). *Case-based reasoning: experiences, lessons, and future directions*, AAAI Press, Menlo Park, Calif.

Lenz, M., Bartsch-Sporl, B., Burkhard, H.-D., and Wess, S. eds. (1998). *Case-based reasoning technology—from foundations to applications*, Springer, Berlin.

Maher, M. L., and Pu, P. eds. (1997). *Issues and applications of case-based reasoning in design*, Lawrence Erlbaum Associates, Mahwah, N.J.

Mitretek. (2000). *QuickZone delay estimation program—user guide, Beta Version 0.91*, ⟨http://www.ops.fhwa.dot.gov/wz/quickz.htm⟩.

Watson, I. (1999). ''Case-based reasoning is a methodology not a technology'' *Knowledge-Based Systems*, 12, 303–308.