Integrating Cases and Models for Prediction in Biological Systems

L. Karl Branting and **John D. Hastings**
Department of Computer Science
University of Wyoming
Laramie, Wyoming 82071-3682
karl@index.uwyo.edu
hastings@uwyo.edu

Jeffrey A. Lockwood
Department of Plant, Soil, and Insect Sciences
University of Wyoming
Laramie, Wyoming 82071
lockwood@uwyo.edu

Abstract

Many complex biological systems are characterized both by incomplete models and limited empirical data. Accurate prediction of the behavior of such systems requires exploitation of multiple, individually incomplete, knowledge sources. *Model-based adaptation* is a technique for integrating case-based reasoning with model-based reasoning to predict the behavior of biological systems. This approach is implemented in CARMA, a system for rangeland grasshopper management advising that implements a process model derived from protocol analysis of human expert problem-solving episodes. CARMA's ability to predict the forage consumption judgments of expert pest managers was empirically compared to that of case-based and model-based reasoning techniques in isolation. This evaluation provided initial confirmation for the hypothesis that an

integration of model-based and case-based reasoning can lead to more accurate predictions than either technique individually.

1 Prediction in Biological Systems

Decision-support in agriculture and natural resources management often requires prediction of the behavior of biological systems. For example, providing advice about the optimal planting time for a crop may require predicting the emergence date of important pests of that crop [PS91]. Similarly, determining the most cost-effective response to a given pest infestation requires predicting crop or forage loss under each available option.

Various approaches to behavioral prediction are possible. In systems for which a precise model exists and accurate values of state variables can be determined, simulation can be used to predict the system's behavior. Alternatively, if there are sufficient historical data, empirical methods such as *case-based reasoning* (CBR) [AP94], decision-tree induction [Qui93], or statistical techniques can be lead to accurate prediction.

Precise models exist for the behavior of many simple physical systems. However, models of agricultural, ecological, and other biological systems are often incomplete, either because a complete state description for such systems cannot be determined or because the number and type of interactions between system elements are poorly understood. Moreover, while historical data often exist for such systems, they are often insufficient for accurate prediction using empirical methods. As illustrated in Figure 1, biological systems often occupy an intermediate point in the continuum between highly analytic domains, such as celestial mechanics and the prediction of artifact behavior, and highly empirical domains, such as sociology [AH92]. In such biological systems, both models and empirical data exist, but neither is *per se* sufficient for accurate prediction. Accurate prediction of the behavior of such systems requires exploitation of multiple, individually incomplete, knowledge sources.

This paper describes the use of *model-based adaptation* as a technique for integrating case-based reasoning with model-based reasoning in domains in which neither technique is individually sufficient for accurate prediction. Under this approach, case-based reasoning is used to find an approximate solution, and model-based reasoning is then used to adapt this approximate solution into a more precise solution. In model-based adaptation, models are used to compensate for insufficient case coverage by extending the range within which cases can be adapted. Conversely, cases compensate for incompleteness in the models by providing a set of reference points with known values.

The next section describes *rangeland pest management*, a task that requires predicting the behavior of a complex biological system, and sets forth a process description of expert problem solving in this domain. Section 3 describes CARMA, a system that implements this process description, and describes how CARMA performs model-based case adaptation. Section 4 describes how CARMA learns match and adaptation weights. An experimental evaluation in which the predictive accuracy of CARMA's model-based adaptation component is compared to that of case-based and model-based reasoning in isolation is set forth in Section 5. This evaluation provides initial confirmation that model-based case adaptation can lead to more accurate simulation of entomologists' predictions than empirical or model-based reasoning alone.

2 Rangeland Pest Management

Rangeland ecosystems typify biological systems having an extensive but incomplete causal theory and limited empirical data. Management tasks for rangelands include optimal stocking rates and grazing systems, water development, wildlife enhancement, noxious weed control, and insect pest management. Each of these management tasks requires evaluating alternative actions by predicting their potential consequences.

The particular rangeland management task of interest to us is pest management. On average, grasshoppers annually consume 21–23% of rangeland forage in the western United States, at an estimated loss of $400 million [HO83]. Rangeland grasshopper infestations can be treated with chemical or biological insecticides, but in many situations the costs of insecticide application exceed the value of the forage saved. Determining the most cost-efficient response to a grasshopper infestation requires predicting the forage savings that would ensue from each response and comparing the savings to the cost of the response itself.

While model-based reasoning can play a role in grasshopper management, there is a general recognition that the interactions affecting grasshopper population dynamics are too poorly understood and too complex to permit precise prediction through numerical simulation [LL91, Pim91, AH92]. However, entomologists and pest managers appear able to provide useful recommendations to ranchers. This indicates that other sources of knowledge can compensate for the absence of a complete model of rangeland ecosystems.

To explicate these knowledge sources and the problem-solving methods employed by experts in applying this knowledge, we performed a protocol analysis of problem solving by several experts in rangeland grasshopper management at the University of Wyoming. For each expert, we transcribed several problem-solving episodes in which the expert responded to a simulated telephone inquiry by a rancher. These "solve-aloud" problem-solving episodes illustrated the elicitation of relevant case facts by the expert, the formation and discrimination among tentative hypotheses, and expert explanations.

The key expert problem-solving step revealed by the protocol analysis was prediction of the proportion of available forage that will be consumed by grasshoppers if no action is taken. Experts appear to perform this predictive step by comparing the current situation to prototypical infestation scenarios. For example, a moderate density of emerging grasshoppers in a cool, wet spring is associated with a low proportion of forage consumption because wet conditions both promote growth of fungal pathogens and increase forage growth. Moreover, cool conditions tend to prolong the early developmental phases[1] during which grasshoppers are most susceptible to pathogens and other mortality factors. In predicting forage consumption by comparing new cases to prototypical scenarios, such as the cool, wet spring prototype, experts appear to be using a form of case-based reasoning.

If a particular new case differs in some ways from a prototype, the expert can perform causal reasoning to predict the effects of the differences. For example, if there is a moderately low density of emerging grasshoppers in a cool, wet spring, an expert will predict low forage consumption because lower density generally means less consumption, and in the prototypical situation low consumption results even from a moderate grasshopper density.

The prototypical infestation scenarios are expressed in terms of abstract features, such as grasshopper species, developmental phases, and density, that are relevant to the expert's model of rangeland ecosystems. In contrast, a rancher's description is almost always in terms of directly observable features, such as the color, size, and behavior of grasshoppers, temperatures, precipitation, *etc.* As a result, determining the most similar infestation scenario requires inferring the relevant abstract features from a set of observations provided by the rancher. Experts exhibit great flexibility in inferring these features. For example, if a rancher is unable to provide the information that discriminates most reliably among grasshopper species (*e.g.*, whether the grasshoppers have slanted faces or a spur on their throats), the expert is able to ask less reliable but easier to answer questions (*e.g.*, "Do the grasshoppers have brightly colored wings or make a clicking sound in flight?").

If the forage consumption will be high enough to lead to forage competition with livestock, the expert determines the interventions that are compatible with local conditions, using knowledge such as that wet conditions preclude the use of malathion, or chemical treatments are precluded by environmental sensitivity. Finally, the expert estimates the relative value of the forage saved in this and future seasons and

[1] During their lifetime, grasshoppers progress through three developmental stages: egg, nymph, and adult. The nymphal stage usually consists of five instars separated by molts. We define the **developmental phases** of a grasshopper's lifecycle to include egg, five nymphal instars, and adult.

the cost of each control measure. The expert then advises the rancher to take most economical action, either applying the most cost-effective control measure or doing nothing. Experts can justify their advice by appeal to an underlying causal model, but seem to use this model only in explaining and adapting the predictions associated with prototypes and not in performing any sort of simulation.

In summary, the protocol analysis indicated that experts in this domain use a highly eclectic reasoning method that includes a form of case-based reasoning for consumption-prediction, rules for inferring case features and acceptable control measures, and causal reasoning for adaptation and explanation. In addition, experts exhibit *opportunistic problem solving* in that they terminate a consultation as soon as the minimum necessary information has been obtained. For example, if the majority of grasshoppers are at too early a stage of development to permit the extent of the infestation to be determined, the rancher is informed that no prediction can be made until later in the season.

3 CARMA: A Rangeland Pest Management Advisory System

We have implemented the problem-solving process described in the previous section in a system termed CARMA (CAse-based Range Management Adviser). The protocol analysis indicated that advice should consist of a treatment recommendation supported by an explanation in terms of causal, economic, and pragmatic factors, including a numerical estimate of the proportion of forage consumed and a cost-benefit analysis of the various treatment options.

The consultation process is as follows:

1. Determine the relevant facts of the infestation case, such as grasshopper species, developmental phases, and density, from information provided by the user. This requires inference rules such as, "if grasshoppers are observed in the spring to

have brightly colored wings or make a clicking sound in flight, then they are bandwinged adults that overwintered as nymphs."

2. Determine whether grasshopper consumption will lead to competition with livestock for available forage.

   (a) Estimate the proportion of available forage that will be consumed by each distinct grasshopper population (*i.e.*, nymphal overwintering, egg overwintering). For each distinct grasshopper population (*i.e.*, subcase):

      i. Determine the prototypical infestation scenario that most closely matches the current subcase. This requires model-based reasoning to assist matching by aligning the developmental phases of the prototypical case and the subcase.

      ii. Adapt the consumption estimate predicted by the prototypical case based on the featural differences between the prototypical and current subcase. This requires model-based reasoning to account for the influence of each feature on consumption.

   (b) Total the forage loss estimates for each subcase to predict the overall proportion of available forage that will be consumed by grasshoppers.

   (c) Compare grasshopper consumption with the proportion of available forage needed by livestock.

3. If there will be competition, determine what possible treatment options should be excluded using rules such as "Wet conditions preclude the use of malathion"; "Environmental sensitivity precludes all chemical treatments."

4. If there are possible treatment options, for each one provide an economic analysis by estimating both the first-year and long-term savings.

   (a) Estimate the first-year savings using model-based reasoning to determine the proportion of forage which would be saved given the efficacy of the treatment type, the developmental phases of the grasshoppers at the time of treatment, and the proportion of lifetime consumption by grasshoppers at each phase.

(b) Estimate the long-term savings using rule-based reasoning to determine if the majority of the grasshoppers will begin laying eggs before treatment can be applied given the developmental distribution of the grasshoppers at the time of treatment. If the majority of grasshoppers will not begin laying eggs, use statistical reasoning to determine the decreased probability of infestation in subsequent years given the Markov transitional probabilities for the infestation location and the effect of the treatment type on beneficial control agents (, *i.e.,* predators and parasites).

To model the ability of human experts for opportunistic problem solving, CARMA terminates a consultation if it discovers any of the following conditions:

- The current date is outside of the season when forage needed for livestock grows.
- The size of the infestation is below the minimum threshold for viability.
- The majority of the grasshoppers overwintered as nymphs. Such grasshoppers divide their consumption between two growing seasons and therefore consume far less during the growing season than grasshoppers overwintering as eggs.
- The majority of the grasshoppers are at such an early developmental phase that the extent of the infestation cannot be predicted with reasonable certainty or at such a late developmental phase that a significant proportion of lifetime forage consumption and egg-laying have already occurred, making treatment futile.

3.1 Determining Relevant Case Features

CARMA begins a consultation by eliciting observations from the user through a series of window-based interface procedures. These observations are used to infer the relevant features of a new case, such as the species, density, and developmental phases of the grasshoppers. CARMA uses multiple levels of rules for inferring each case feature, ordered by the certainty or the accuracy of each rule. The rules are applied in succession until either the user can provide the necessary information or a default value is chosen. For example, if the value of the case feature "total number of grasshoppers per square yard" is unknown to the user, CARMA instructs the user to estimate the number of grasshoppers that would be present in 18 square-foot circles. If the user can't provide this information, the system attempts to infer this feature using a rule that grasshopper density is equal to 1.5 times the number of grasshoppers seen hopping away with each step taken by the user in the field. Otherwise, the value defaults to the statewide historic average of four grasshoppers per square yard. By applying rules in the order of their certainty, CARMA reasons with the best information available.

A typical interface window for determining the observed grasshopper type distribution appears in Figure 2. It includes the options "Why" for describing why this information is important to the consultation, "Help" for advising the user about the various window features and their operations, "How To" to explain the proper procedure for gathering the required information, "Not sure" to trigger the selection of an alternative rule for inferring the feature, and "OK" to indicate that the user has chosen an answer. "Display planthopper" shows a small insect that the user should distinguish from a grasshopper. Figure 3 shows an input window that asks the user to provide the infestation location by clicking on a map of Wyoming's major roads, towns, and county borders. CARMA uses this location to retrieve the historical values for the site including infestation history, range value, temperature, and precipitation.

Since a complete case specification is not always required for useful advice, CARMA fills in the facts of a new case opportunistically. This means that CARMA asks the user for information only when the corresponding case feature is required for the reasoning process to continue. At the earliest point at which a decision can be made, the case-feature inference process halts, advice is given, and the consultation is completed. This minimizes the amount of input required for CARMA to make a decision, thereby accelerating consultations. For example, if the date and location of an infestation indicate that it is too early to assess the severity of a grasshopper infestation, CARMA advises the user to rerun the consultation at a later time without prompting for further information.

3.2 Case Matching

The protocol analysis indicated that pest managers estimate forage consumption by comparing new cases to prototypical infestation scenarios. These prototypical cases differ from conventional cases in two important respects. First, the prototypical cases are not expressed in terms of observable features (*e.g.*, "Whenever I take a step, I see four grasshoppers with brightly colored wings fly"), but rather in terms of abstract derived features (*e.g.*, "Approximately six nymphal overwintering grasshoppers in the adult phase per square yard"). Second, the prototypical cases are extended in time, representing the history of a particular grasshopper population over its lifespan. Each prototypical case is therefore represented by a "snapshot" at a particular, representative point in time selected by the entomologist. In general, this representative point is one at which the grasshoppers are at a developmental phases in which treatment is feasible. An example prototypical case appears as Case4 in Table 1.

A tract of rangeland almost invariably contains multiple grasshopper species, which may differ widely in consumption characteristics. In particular, grasshoppers that spend the winter as nymphs consume far less during the growing season than grasshoppers overwintering as eggs. CARMA therefore partitions the overall population of a case into subcases according to life history (*i.e.*, overwintering as nymphs or eggs). The overall grasshopper population is initially divided into three observed categories: bandwinged (*i.e.*, grasshoppers having brightly-colored wings or make a clicking sound in flight); forb or mixed grass/forb feeders (*i.e.*, grasshoppers having a round head with a spur throat); and grass feeders (*i.e.*, grasshoppers having a slanted face or pointed head, or a round head with no spur throat). If the grasshoppers are part of the bandwinged category, CARMA concludes that the grasshopper population is nymphal-overwintering. Otherwise, the population is determined to be egg-overwintering. For example, the new case set forth in Table 1 is split into two subcases, SubcaseA and SubcaseB, based on overwintering type.

To predict the forage loss of a subcase, CARMA first retrieves all prototypical cases whose life history (*i.e.*, overwintering type) matches that of the subcase. The weighted sum of featural differences between each prototypical case and the new subcase is calculated to determine the most similar prototypical case. Match weights are determining from the *mutual information gain* between case features and qualitative consumption categories in a given set of training cases, since recent research has indicated that this is often the most accurate measure of featural importance for matching [WD95]. Separate match weights are computed for each grasshopper overwintering type for the seven case features: precipitation, temperature, range value, infestation history, average developmental phase, density, and feeding type. Quantitative features, such as density, are converted to qualitative values for computation of mutual information gain, since small quantitative variations seemed to have little effect on matching. The difference between two individual feature values is determined by finding the difference between the positions of the values in an ordered qualitative feature value list. For example, range value can equal one of the qualitative values in the ordered set {low, low-moderate, moderate, high-moderate, and high}, so that the matching feature difference between low and high, the maximum possible difference, is four. The forage loss prediction associated with the best matching prototypical case is then adapted to apply to the current subcase.

3.3 Model-Based Adaptation

The assumption underlying model-based adaptation is that the causal models associated with a biological or other partially understood systems may be accurate in the neighborhood of a case, even if the models are not *per se* sufficient for accurate prediction throughout the entire feature space. CARMA uses three specific forms of model-based adaptation: temporal project; featural adaptation; and critical period adaptation. The details of these adaptation methods reflect the particular causal models associated with rangeland ecosystems. However, we believe that the general approach of performing simulation or other model-based reasoning to adapt a case to apply to new cases in its neighborhood in feature space has applicability to a wide range of biological systems.

3.3.1 Temporal Projection

Since prototypical cases are extended in time but are represented at a particular moment, CARMA must project the best matching prototypical case forward or backwards in time to align its average developmental phase with that of the new subcase. This requires using a model to simulate grasshopper attrition, which depends on developmental phase, precipitation, and developmental rate (which in turn depends on temperature) throughout the interval of the projection. CARMA assumes that the grasshoppers within a developmental phase are evenly distributed throughout the phase. Therefore, CARMA breaks the distribution into daily populations, projects the populations the required number of days (adjusting the density each day based on attrition), then regroups the daily populations into their new developmental phases. Attrition rates are adjusted by scalars (one scalar for precipitation = wet, and another for precipitation = non-wet) that are learned via the algorithms described in Section 4. A graphic example of temporal projection appears in Figure 4.

For example, the prototypical case that best matches SubcaseA is Case4, shown in Table 1. Because the developmental phase of Case4 before projection is earlier than that of SubcaseA, the population in Case4 must be projected forward in time in order for it to be at the same stage of development as the population in SubcaseA. Projection forward in time causes grasshoppers to be removed from the population due to attrition (*i.e.*, 27.0 grasshoppers per square yard before projection to 24.0 grasshoppers per square yard after projection). Temporal projection aligns developmental phases but not necessarily dates.

3.3.2 Featural Adaptation

The forage loss predicted by the best matching prototypical case, FL(PC), is modified to account for any featural differences between it and the subcase, based on the influence of each of the n features on consumption as represented by a list of featural adaptation weights $\bar{A} = (A_1, \ldots, A_n)$. Thus, the predicted forage loss for the new subcase, $FL(NC)$, is determined as follows:

$$FL(NC) = FL(PC) + \sum_{i=1}^{n} A_i * QFD(i)$$

where $QFD(i)$ is the quantitative difference for feature i between the new subcase and prototypical case. For example, a lower temperature value means lower forage losses, because lower temperatures tend to slow development, increasing grasshopper attrition. Thus, the forage loss estimate predicted by Case4—60%—must be adapted downward somewhat to account for the fact that temperatures in SubcaseA (cool) are lower than in Case4 (normal). In determining the quantitative feature difference between the new subcase and the prototypical case for qualitative features such as temperature, CARMA computes a simple difference:

$$Q(NC, i) - Q(PC, i)$$

where $Q(NC, i)$ and $Q(PC, i)$ are the quantitative values for feature i in the new subcase and prototypical case, respectively. For quantitative features such as density, proportion of lifetime consumption in the critical period, and total area infested, a proportional difference is used:

$$\frac{Q(NC, i) - Q(PC, i)}{Q(PC, i)}$$

Adaptation weights are set using a hill-climbing algorithm that optimizes CARMA's predictive accuracy on training instances (discussed in Section 4). The weights used in featural adaptation can be viewed as a linear approximation of the function from derived case features to consumption amounts in the neighborhood of each prototypical case.

3.3.3 Critical Period Adaptation

Grasshopper consumption is most damaging if it occurs during the *critical forage growing period*, *i.e.,* the portion of the growing season during which forage losses caused by grasshoppers cannot be fully replaced by forage growth. The forage loss predicted by a prototypical case must be adapted if the proportion of the lifespan of the grasshoppers overlapping the critical period in the new case differs from that in the prototypical case. This process, termed *critical period adaptation*, requires determining the proportion of lifetime consumption occurring in the critical period based on the developmental phases of the new and prototypical cases that fall within the critical period and the proportion of lifetime consumption occurring in these developmental phases. The forage loss estimate is then adjusted based on the featural adaptation weight for the critical period and the difference in the proportion of lifetime consumption in the critical period between the new case and prototypical case.

A graphic example of critical period adaptation is hown in Figure 5. Because grasshopper development in SubcaseA is ahead of that in Case4 (SubcaseA's developmental phase on June 14 corresponds to Case4's developmental phase on June 15), CARMA determines that Case4 applies to more of the critical period than SubcaseA because it will only reach Day 1 of developmental phase 3 by the beginning of the critical period (June 17), while SubcaseA will already reach Day 8 of developmental phase 3. CARMA uses a model of grasshoppers' rate of consumption at each developmental phase to calculate the proportion of lifetime consumption occurring after the beginning of the critical period and before the end of the critical period. For example, only 86.0% of SubcaseA's consumption occurs during the critical period, whereas 92.7% of Case4's consumption occurs within this period. The quantitative feature difference for critical period adaptation is computed as a proportional difference, therefore CARMA adjusts the initial consumption estimate by $(86.0 - 92.7) / 92.7 = -0.072$ multiplied by the adaptation feature weight for critical period.

In summary, CARMA uses a model of grasshopper developmental phases, consumption, and attrition, knowledge concerning the relative contribution of case features to consumption, and a model of a rangeland's critical forage growth period in adapting the cases in its library.

3.4 Forage Loss Estimation

After adaptation, the consumption predictions for each subcase (*i.e.*, populations of grasshoppers with distinct feeding patterns) are summed to produce an overall consumption estimate. In the given case, the sum of predicted consumption of the two subcases is 90% (86.5 + 3.4). Because of variability resulting from the imprecise nature of rangeland ecosystems, this prediction is converted to the qualitative range, high, meaning that approximately 60 to 100% of the available forage will be lost. An interface window explaining estimated forage loss is shown in Figure 6. It gives both aggravating and mitigating factors (*i.e.*, factors tending to increase vs. reduce the forage loss estimate).

The natural language explanation is produced using conventional template instantiation techniques. First, the explanation generator creates the natural language representation of pertinent qualitative feature values using simple lookup tables (*e.g.*, the text string for feature value high-mod is "moderately high"). The text strings are then combined with the explanation template. For example, the template for the first sentence in the forage loss explanation is:

> < "From the information you have provided, it is estimated that the grasshoppers will consume a " qualitative-forage-loss-string " percentage of the forage available for the year or approximately " quantitative-forage-loss-range-string "%.">

If the proportion of available forage that will be lost to grasshoppers and the proportion needed for livestock (and wildlife) exceeds 100% of the forage available, CARMA concludes that competition will occur. In this example, competition is possible and the consultation should continue if the proportion of available forage needed by livestock is greater than 40%. For example, if forage need is 60%, the expected year-long competition should range from 0% (*i.e.*, (40 + 60) − 100) to 20% (*i.e.*, (60 + 60) − 100). A typical interface window explaining estimated forage competition is shown in Figure 7.

3.5 Determining Treatment Options

If there will be competition, CARMA applies a set of rules to determine what possible treatment options are excluded by the conditions of the case. Some of the information necessary for determining exclusion is already known from the case features (*e.g.*, the presence of grasshoppers in the first nymphal instar suggests an ongoing hatch, thereby excluding malathion and carbaryl bait from consideration). Other conditions must be determined from further user input (*e.g.*, "Will it be hot at the time of treatment?" If so, exclude malathion). An interface window explaining the selection of acceptable treatments appears in Figure 8. The explanation includes the rules that were used to exclude treatments. This explanation is also derived using standard template-instantiation techniques.

3.6 Treatment Recommendation

For each possible treatment option, CARMA provides estimates of the reduced probability of future reinfestation and current-year and long-term savings. From the estimated savings, CARMA recommends the treatment or treatments that are most economical. A typical treatment recommendation window including estimates of future reinfestation and economic savings appears in Figure 9. Note that this analysis includes "no treatment" as an option.

3.6.1 Reduced Probabilities of Future Reinfestation.

CARMA uses Markov transitional probabilities for the infestation location (derived from historical infestation history data collected by the USDA and synthesized by the University of Wyoming Entomology Section) to calculate for each treatment type the total reduced probability of future reinfestation.

CARMA first determines whether the grasshoppers will begin laying eggs before the treatment date. If the developmental distribution of the grasshoppers at treatment is dominated by adults, CARMA determines that too many eggs will already be laid, and no reduction in the probability of future reinfestation will result from treatment because eggs are not affected by treatment. If few eggs will have been laid, CARMA calculates the yearly reinfestation probabilities for each treatment type based on the historical Markov transitional probabilities for as many years as the probability of infestation with treatment is significantly lower that the probability of infestation without treatment (*i.e.,* until the benefits of treatment have ended). The total reduced probability of future reinfestation for each treatment is calculated by summing each yearly difference between the probabilities of infestation without and with treatment.

Because the number of grasshoppers that may emerge in future years is often not directly proportional to the number of eggs laid the current year (*e.g.,* under ideal conditions, grasshoppers are capable of expanding from a low population on year to a very high population the next), transitional probabilities are adjusted only slightly based on the efficacy of treatments in reducing the number of eggs laid. The transitional probabilities are reduced further for those treatments capable of preserving beneficials. For example, treatments such as carbaryl bait are designed to be consumed specifically by grasshoppers and are therefore unlikely to affect biological control agents such as birds and insects. Conversely, sprays such as malathion blanket an entire area and hurt beneficials indiscriminately. A greater reduction in the transitional probabilities is made for treated infestations whose total area is quite large, because treatment will tend to reduce the chance that grasshoppers from previously untreated areas will migrate into the treated area.

3.6.2 Economic Analysis

For each possible treatment option, CARMA provides estimates of current-year and long-term savings. Each analysis involves a range that indicates best to worst case estimates (negative values indicate a loss). A typical interface window explaining the savings calculations appears in Figure 10.

Current-year Savings. For each possible treatment option, CARMA estimates the current-year savings as the difference between the value of forage in competition saved by treating and the treatment cost. CARMA first computes the amount of pre-treatment forage loss. This is done by projecting the developmental distribution of each subcase forward to the user-provided treatment date (often a week or more from the current date). In a manner similar to determining the percentage of lifetime consumption occurring within the critical period, CARMA applies a model of grasshoppers' rate of consumption at each developmental phase to each subcase to calculate the proportion of lifetime consumption occurring before the treatment date. This proportion is used to scale the year-long forage loss estimate, resulting in the pre-treatment loss. The pre-treatment forage loss estimates for each subcase are summed to produce the total pre-treatment forage loss. Next, CARMA estimates the amount of post-treatment forage loss without treatment by subtracting pre-treatment forage loss from total forage loss. For example, if total forage loss is estimated to be 60 to 100%, and pre-treatment forage loss is estimated to be 2.0 to 3.3%, then the post-treatment forage loss will be 58.0 to 96.7%.

For each option, CARMA estimates the amount of post-treatment forage loss with treatment according to the expected efficacy of the treatment and the post-treatment forage loss without treatment. For example, the insecticide carbaryl formulated as a baitbait is usually 65 to 80% effective. If the estimated post-treatment forage loss without treatment is 58.0 to 96.7%, then at best carbaryl bait should prevent 80% of the 58.0% loss, and at worst prevent 65% of the 96.7% loss, resulting in a 11.6 to 33.9% post-treatment forage loss.

CARMA calculates the year-long forage loss for each option by summing pre- and post-treatment forage loss. Year-long competition resulting from a treatment option is calculated by comparing year-long forage loss resulting from the option and forage need. The proportion of forage in competition saved is simply the proportion of forage in competition without treatment minus the proportion of forage in competition with treatment. For example, if pre-treatment forage loss is 2.0 to 3.3% and post-treatment forage loss is 11.6 to 33.9%, the year-long forage loss for the option is 13.6 to 37.2%. Given a forage need of 60%, the year-long competition with treatment ranges from $(13.6 + 60) - 100 = -26.4$ to $(37.2 + 60) - 100 = -2.8$, which is less than zero, thereby preventing competition. If the year-long forage in competition without treatment is 20 to 60%, and the treatment option will result in no competition, then the expected forage in competition saved by treating is 20 to 60%.

With the per-unit forage value and range value estimates provided by the user, CARMA estimates the current-year savings for an option to be the value of forage in competition that is saved minus the cost of the treatment. In this example, the per-unit forage value is \$30/AUM (*i.e.*, an animal unit month: the amount of forage necessary to support a cow and calf for one month) and the estimated range value (or productivity) is 6-10 acres/AUM. Therefore, the current-year savings ranges from:

$$20\% \times \frac{\$30}{AUM} \times \frac{AUM}{10acres} = \frac{\$0.60}{acre}$$

to

$$60\% \times \frac{\$30}{AUM} \times \frac{AUM}{6acres} = \frac{\$3.00}{acre}$$

Long-term Savings. CARMA calculates the savings for future years for each treatment type as the value of year-long (*i.e.*, total) forage in competition without treatment (taken from the first year calculations) times the total reduced probabilities of future reinfestation. Based on the current-year savings, CARMA recommends the treatment that is estimated to save the most under a worst-case scenario and the treatment that is estimated to save the most under a best-case scenario. Usually, the worst and best scenarios produce the same recommended treatment. Following the treatment recommendation, the consultation is complete.

4 Learning Match and Adaptation Weights

CARMA uses two sets of weights in case-based reasoning: match weights (used in the assessment of similarity between cases) and featural adaptation weights (used to adapt the consumption predicted by the best matching prototypical case in light of any featural differences between it and the subcase). General domain knowledge, such as the identifying characteristics and developmental phases of grasshoppers, can be provided by the domain expert. By contrast, match and featural adaptation weights must be acquired by the system itself.

As indicated above, match weights are set by determining the mutual information gain between case features and qualitative consumption categories in a given set of training cases.

Featural adaptation weights are set by a hill-climbing algorithm, `AdaptWeights`, that incrementally varies adaptation weights $\bar{A}$ to minimize the *root-mean-squared error* (RMSE):

$$\sqrt{1/n \sum_{i=1}^{n} [\texttt{PFL}(C_i, P, M, \bar{A}) - \texttt{ExpertPred}(C_i)]^2}$$

for prototypical case library P and match weights M, where $\texttt{PFL}(C_i,\ P,\ M,\ \bar{A})$ is CARMA's predicted forage loss and $\texttt{ExpertPred}(C_i)$ is the expert's prediction of consumption for each training case C_i. The algorithm for `AdaptWeights` is as follows:

function AdaptWeights(T, P, M)

$I \leftarrow$ initial increment

$D_{min} \leftarrow$ minimum improvement threshold

$I_{min} \leftarrow$ minimum increment threshold

$\bar{A} \leftarrow$ initial list of global adaptation weights

$D' \leftarrow$ RMSE(T,P,M,$\bar{A}$)

$D \leftarrow \infty$

loop until $(I < I_{min})$ **do**

  loop until $(|D' - D| < D_{min})$ **do**

$D \leftarrow D'$
$\delta \leftarrow$ the change to an element of $\bar{A}$ by I for which RMSE(T,P,M,$\delta(\bar{A})$) is least
$D' \leftarrow$ RMSE(T,P,M,$\delta(\bar{A})$)
if $(D' < D)$ **then** $\bar{A} \leftarrow \delta(\bar{A})$
else $D' \leftarrow D$
$I \leftarrow I/2$
return $\bar{A}$

Separate adaptation weights are computed for each grasshopper overwintering type for the same eight case features: precipitation, temperature, range value, infestation history, average developmental phase, density, feeding type, proportion of lifetime consumption in the critical period, and total area infested. In computing the featural adaptation weights, qualitative case features (such as precipitation = Dry) are converted into quantitative values based on the position of the value in an ordered qualitative feature value list. An adaptation feature difference is computed as the difference between the quantitative feature values of the two cases. The consumption prediction of the matching prototypical case is adjusted by the sum of the adaptation feature differences multiplied by the adaptation weights for each feature. CARMA can learn featural adaptation weights in either of two modes: *global*, in which a single set of weights are acquired for the entire entire case library or *case-specific*, in which separate weights are acquired for each prototypical case.

5 Evaluating Model-Based Adaptation

The design of CARMA's forage consumption component was based on the hypothesis that an integration of model-based and case-based reasoning can lead to more accurate forage consumption predictions than the use of either technique individually. This hypothesis is based on the observation that neither the causal model nor the empirical data available for rangelands are individually sufficient for accurate prediction. To test this hypothesis, we separated CARMA's empirical and model-based knowledge components, tested each in isolation, and compared the results to the performance of the full CARMA system under both global and case-specific adaptation weight modes.

The evaluation was complicated by the absence of empirical data against which to measure CARMA's predictions. We therefore turned to expert human judgments as an external standard. To obtain a representative sample of expert opinions, we sent questionnaires to 20 entomologists (including pest managers) recognized for their work in the area of grasshopper management and ecology. Each expert received 10 cases randomly selected from a complete set of 20 hypothetical cases set in northern Wyoming. The descriptions of the 20 cases contained at least as much information as is typically available to an entomologist from a rancher seeking advice. The questionnaire asked the expert to predict quantitative forage loss and the most appropriate course of action. A total of 15 recipients of the questionnaire responded. There was a very wide variation (from 25 to 90%) in consumption predictions of the respondents over the set of 20 cases. However, there appeared to be a much higher degree of consistency among the eight experts from Wyoming, so in the experiments described below we restricted our attention to the eight sets of responses from Wyoming experts.

5.1 Experimental Design

Each predictive method was tested using a series of leave-one-out tests in which a set of cases (S) from a single expert was split into one test case (C) and one training set (S − C). The methods were trained on the forage loss predictions of the training set and tested on the test case. This method was repeated for each case within the set (S). The forage loss predictions (between 0% and 100%) represent the proportion of available forage that would otherwise be available for livestock, but will instead be consumed by grasshoppers. CARMA was tested using a protocol under which each set of training cases was used as CARMA's library of prototypical cases. This protocol is implemented in LeaveOneOutSpecificTest and LeaveOneOutGlobalTest, which perform the leave-one-out tests for the specific and global adaptation weighting schemes, respectively. Both procedures call AdaptWeights, the hill-climbing algorithm described above. LeaveOneOutSpecificTest calls AdaptWeights with a prototypical case library containing only one case.

```
function LeaveOneOutSpecificTest(T)
    for each case Ci ϵ T do
        P ← T − Ci          ;prototypical cases
        M ← global match weights for set P
                according to info. gain
        for each prototypical case Pj ϵ P do
            T ← P − Pj          ;training set
            Pj(A) ← AdaptWeights(T, {Pj}, M)
        Di ← (PredictForageLoss(Ci, P, M)
                − ExpertPred(Ci))^2
    return ( √Avg(D))
```

```
function LeaveOneOutGlobalTest(T)
    for each case Ci ϵ T do
        P ← T − Ci          ;prototypical cases
        M ← global match weights for set P
                according to info. gain
        G ← AdaptWeights(P, P, M)
        Di ← (PredictForageLoss(Ci, P, M, G)
                − ExpertPred(Ci))^2
    return ( √Avg(D))
```

CARMA's empirical component was evaluated by performing leave-one-out-tests for CARMA's forage consumption module with all model-based adaptation disabled. CARMA's forage consumption module with model-based adaptation disabled is termed *factored nearest-neighbor prediction* (factored-NN), because under this approach prediction is based simply on the sum of nearest neighbor predictions for each subcase. Two other empirical methods were evaluated as well: decision-tree induction using ID3[2][Qui93] and linear approximation using QR factorization [Hag88] to find a least-squares fit to the feature values and associated predictions of the training cases.

The predictive ability of CARMA's model-based component in isolation was evaluated by developing a numerical simulation based on CARMA's model of rangeland ecology. This simulation required explicit representation of two forms of knowledge implicit in CARMA's cases: the forage per acre based on the range value of the location, and the forage typically eaten per day per grasshopper for each distinct grasshopper overwintering type and developmental phase. The steps of the numerical simulation are as follows:

1. Project each grasshopper population back to beginning of the growing season.

2. Simulate the density and developmental phases for each overwintering type through the end of the critical period growth season based on the precipitation and temperature given in the case.

3. Calculate the forage eaten per day per acre based on the grasshopper density per acre and the forage eaten per day per grasshopper for each overwintering type and developmental phase as affected by temperature.

4. Convert the total forage consumed to the proportion of available forage consumed based on the forage per acre.

The effect of temperature on consumption (as a result of changing metabolic rates) was represented by multiplying a coefficient (determined from a lookup table indexed by temperature) by the forage eaten per day per grasshopper for each overwintering

[2]ID3 classified cases into 10 qualitative consumption categories representing the midpoints (5, 10, 15, ... , 95) of 10 equally sized qualitative ranges. ID3's error was measured by the difference between the midpoint of each predicted qualitative category and the expected quantitative consumption value.

type. The numerical simulation was trained by hill-climbing on temperature-based coefficients to maximize the predictive accuracy of the training cases.

5.2 Results

The accuracy of each approach was tested using leave-one-out testing for each of the eight Wyoming Expert Sets and for a data set consisting of the median of the predictions of the Wyoming experts on each case. The results, which appear in Table 2, include the root-mean-squared error for each of the methods.

The results of the integration experiment provide initial confirmation for the hypothesis that integrating model-based and case-based reasoning through model-based adaptation leads to more accurate forage consumption predictions than the use of either technique individually. The lowest root-mean-squared error rate was obtained by CARMA-specific. On the Wyoming Expert Sets, the root-mean-squared error rate was 13.3 for CARMA-specific and 14.2% for CARMA-global. The root-mean-squared error rate was higher both for the empirical approaches—21.1% for factored-NN, 34.9% for ID3, and 25.6% for linear approximation—and for the purely model-based approach—29.6%. CARMA-specific and CARMA-global were also more accurate than the alternative methods on the Wyoming median set, although linear approximation was only slightly less accurate. The initial confirmation of the hypothesis that integrating model-based and case-based reasoning through model-based adaptation leads to more accurate forage consumption predictions than the use of either technique individually is tentative because the low level of agreement among experts and the absence of any external standard give rise to uncertainty about what constitutes a correct prediction. However, this validation problem appears to be an inherent property of biological domains such as rangeland pest management.

Consumption prediction can be viewed as approximating a function from derived case features to consumption predictions (a consumption function). Prototypical cases constitute representative points in feature space for which function values are known. The prototypical cases can be used to induce a representation of the function as a decision tree (*e.g.*, ID3) or a numerical function (*e.g.*, linear approximation). The poor performance of ID3 and linear approximation suggests that the biases of these inductive methods are poorly suited to the consumption prediction task. The high performance of linear approximation on the Wyoming median set (11.9%) suggests that taking the median of the predictions for the expert sets causes the complex consumption function curve to be drastically flattened, with the result that it is much more easily predicted by linear approximation.

Numerical simulation can be used to derive individual values for the function. However, the incompleteness of available models of rangeland ecology limits the accuracy of this approach. A pure nearest-neighbor approach implicitly assumes that the consumption function is constant in the neighborhood of prototypical cases. CARMA's model-based adaptation approach uses a model of rangeland ecology to approximate the consumption function in the neighborhood of individual prototypical cases. For example, projection consists of simulation through the temporal interval necessary to align the developmental phases of two cases. Although the model may be insufficient in itself for accurate consumption prediction, it may greatly improve the accuracy of nearest-neighbor prediction.

In summary, the tests of CARMA's forage consumption prediction component provide an initial confirmation of the hypothesis that integrating model-based and case-based reasoning can lead to more accurate forage consumption predictions than the use of either technique individually.

6 Status

On June 17, 1996, CARMA was distributed to the University of Wyoming Cooperative Extension offices and Weed and Pest District Offices in each of the 23 Wyoming counties. The fielded version consists of CARMA-specific using a case library consisting of the Wyoming median set. CARMA is available free of charge for non-commercial purposes and can be down-loaded from http://ai.uwyo.edu/~karl/carma

CARMA is implemented in Allegro CL/PC and runs under Windows 3.1 or Windows-95 on 486 or higher processors with a minimum of 8MB of RAM and 21MB of swap space.

7 Related Work

Several previous research projects have investigated the benefits of integrating case-based reasoning with model-based reasoning. However, these projects have generally assumed the existence of a correct and complete causal model. For example, CASEY [Kot88] performed diagnosis using model-based reasoning to assist both case matching and case adaptation. However, CASEY presupposed both the existence of a complete causal theory of heart disease and complete explanations of each case in terms of that theory. Goel's use of device models to adapt design cases also presupposed that the device models are complete and correct [Goe91]. Similarly, Rajamoney and Lee's *prototype-based reasoning* [RL91] presupposed a complete and correct (though not necessarily tractable) causal model.

Feret and Glascow [FG93] described an alterative approach under which model-based reasoning is used for "structural isolation" (*i.e.*, identification of the structural components of a device that probably give rise to the symptoms of a fault). Cases are indexed by these tentative diagnoses, which are then refined using case-based reasoning. This approach, while appropriate for diagnosis, is ill-suited for behavioral prediction in the absence of faults.

CARMA's technique of model-based matching and adaptation represents an alternative approach to integrating CBR and MBR in domains characterized by an incomplete causal model.

8 Conclusion

This paper has described a technique for integrating case-based reasoning with model-based reasoning to predict the behavior of biological systems characterized both by incomplete models and insufficient empirical data for accurate induction. This technique is implemented in CARMA, a system for rangeland pest management advising. An empirical evaluation provided confirmation of the hypothesis that integrating model-based and case-based reasoning through model-based adaptation can lead to more accurate forage consumption predictions than the use of either technique individually. We believe that the approach to model-based adaptation embodied in CARMA is appropriate for other domains in which empirical and model-based knowledge are individually insufficient for accurate prediction, such as predictive tasks involving biological, ecological, and other complex natural systems.

Acknowledgements

This research was supported in part by a grant from the University of Wyoming College of Agriculture, a Faculty Grant-in-Aid from the University of Wyoming Office of Research, and by NSF Career Grant IRI-9502152.

References

[AH92] T. F. H. Allen and T. W. Hoekstra. *Toward a Unified Ecology.* Columbia University Press, New York, NY, 1992.

[AP94] A. Aamodt and E. Plaza. Case-based reasoning; foundational issues, methodological variations, and system approaches. *AI Communication*, 7(1):39–59, March 1994.

[FG93] M. P. Feret and J. I. Glascow. Hybrid case-based reasoning for the diagnosis of complex devices. In *Proceedings of Eleventh National Conference on*

Artificial Intelligence, pages 168–175, Washington, D.C., July 11–15 1993. AAAI Press/MIT Press.

[Goe91] A. Goel. A model-based approach to case adaptation. In *Thirteenth Annual Conference of the Cognitive Science Society*, pages 143–148, 1991.

[Hag88] W. Hager. *Applied Numerical Linear Algebra*. Prentice Hall, 1988.

[HO83] G. B. Hewitt and J. A. Onsager. Control of grasshoppers on rangeland in the united states: a perspective. *Journal of Range Management*, 36:202–207, 1983.

[Kot88] P. Koton. *Using Experience in Learning and Problem Solving*. PhD thesis, Massachusetts Institute of Technology, 1988. Department of Electrical Engineering and Computer Science.

[LL91] J. Lockwood and D. Lockwood. Rangeland grasshopper (Orthoptera: Acrididae) population dynamics: Insights from catastrophe theory. *Environmental Entomology*, 20:970–980, 1991.

[Pim91] S. L. Pimm. *The Balance of Nature: Ecological Issues in the Conservation of Species and Communities*. University of Chicago Press, Chicago, 1991.

[PS91] R. Plant and N. Stone. *Knowledge-Based Systems in Agriculture*. McGraw-Hill, New York, 1991.

[Qui93] R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Mateo, CA, 1993.

[RL91] S. Rajamoney and H. Lee. Prototype-based reasoning: An integrated approach to solving large novel problems. In *Proceedings of Ninth National Conference on Artificial Intelligence*, Anaheim, July 14–19 1991. AAAI Press/MIT Press.

[WD95] D. Wettschereck and T. Dietterich. An experimental comparison of the nearest-neighbor and nearest-hyperrectangle algorithms. *Machine Learning*, 19(1):5–28, 1995.

Biographical Sketches

Karl Branting is assistant professor of computer science at the University of Wyoming. He received a B.A. in philosophy from the University of Colorado, a J.D. from Georgetown University Law Center, and a Ph.D. in computer science from the University of Texas at Austin. His research interests include case-based reasoning, machine learning, cognitive modeling, and applications of artificial intelligence to natural resources management, ecology, and law.

Jeff Lockwood is professor of entomology at the University of Wyoming. He received a B.S. in biology from New Mexico Tech and a Ph.D. in entomology from the Louisian State University. His research interests include insect ecology, with particular attention to grasshoppers and their role in rangeland ecosystem processes.

John Hastings received a B.S., M.S., and Ph.D. in computer science from the University of Wyoming. His 1996 dissertation, "A Mixed Paradigm Reasoning Approach To Problem-Solving In Incomplete Causal-Theory Domains", described the development, implementation, and testing of CARMA.

	Case4	New case		Case4
		SubcaseA	SubcaseB	after projection
Overwintering type	nymph	nymph	egg	egg
Feeding types	grass 10% mixed 90%	grass 50% mixed 50%	grass 100%	grass 10% mixed 90%
Average phase	2.0	3.0	7.0	3.0
Density	27.0	36.0	4.0	24.0
Proportion of lifetime consumption in critical period	92.7	86.0	12.4	92.7
Date	June 8	June 14		June 15
Precipitation	normal	dry		normal
Temperatures	normal	cool		normal
Infest. history	high	high		high
Range value	low	moderately-high		low
Total area infested	12000	9800		12000
Forage loss	60% (high)	?		60% (high)

Table 1: Case examples.

	CARMA		Empirical Only			Model-Based Only
	Specific weights	Global weights	Factored-NN	ID3	Linear appr.	Numerical simulation
Wyoming expert sets	13.3	14.2	21.1	34.9	25.6	29.6
Wyoming median set	9.7	10.0	22.8	35.2	11.9	28.8

Table 2: Root-mean-squared error (in %) for leave-one-out-test results.