The Thirty-Second International Florida
Artificial Intelligence Research Society Conference (FLAIRS-32)

Exploiting Markov Random Fields to Enhance Retrieval in Case-Based Reasoning

Luigi Portinale
Computer Science Institute, DiSIT
Università del Piemonte Orientale
Alessandria, Italy 15121
luigi.portinale@uniupo.it

Abstract

The similarity assumption in Case-Based Reasoning (similar problems have similar solutions) has been questioned by several researchers. If knowledge about the adaptability of solutions is available, it can be exploited in order to guide retrieval. Several approaches have been proposed in this context, often assuming a similarity or cost measure defined over the solution space. In this paper, we propose a novel approach where the adaptability of the solutions is captured inside a metric Markov Random Field (MRF). Each case is represented as a node in the MRF, and edges connect cases whose solutions are close in the solution space. States of the nodes represent the adaptability effort with respect to the query. Potentals are defined to enforce connected nodes to share the same state; this models the fact that cases having similar solutions should have the same adaptability effort with respect to the query. The main goal is to enlarge the set of potentially adaptable cases that are retrieved (the recall) without significantly sacrificing the precision of retrieval. We will report on some experiments concerning a retrieval architecture where a simple kNN retrieval is followed by a further retrieval step based on MRF inference.

Keywords: Case-Based Reasoning, Adaptation Guided Retrieval, Markov Random Fields.

1 Introduction

The main postulate of Case-Based Reasoning (CBR) is that "similar problems have similar solution(s)" (*similarity assumption*). The more valid the similarity assumption, the more efficient the CBR process is, since the retrieved solutions are more similar to the (unknown) solution to the query.

The most common retrieval strategy is based on *k-Nearest Neighbor* (kNN) algorithms, returning the solutions of the k cases stored in the library that have the most similar description with respect to the query; let us call it *structural similarity* in contrast to *solution similarity*. The postulate of similarity assumption has been questioned by several researchers, since there are problems (and applications) where it is unsafe to rely on it (Smyth and Keane 1998; Stahl and Schmitt 2002; Abdel-Aziz, Strickert, and Hüllermeier 2014). In particular, when structural similarity is not a real proxy for solution similarity, new retrieval techniques should be devised, in order to guarantee that retrieved solutions are likely to be useful (i.e. reusable or at least revisable). Different solutions have been devised to address this problem: the introduction of specific or task dependent adaptation knowledge into the retrieval step in adaptation-guided retrieval (Smyth and Keane 1998; Portinale, Torasso, and Magro 1997; Diaz-Agudo, Gervas, and Gonzales-Calero 2003), the modeling of solution preferences in preference-based CBR (Abdel-Aziz, Strickert, and Hüllermeier 2014), the learning of a utility-oriented similarity measure minimizing the discrepancies between the similarity values and the desired utility scores (Xiong and Funk 2006).

In this paper, we describe a novel technique where standard kNN retrieval is improved through an adaptation guided inference process based on Markov Random Field (MRF), inference (Murphy 2013). In the following, we first introduce some basic notions concerning MRFs and metric MRFs in particular. We then describe the characterization of case adaptability we rely on, and we discuss a framework for case retrieval based on inference on an MRF, capturing the relevant adaptation knowledge. A retrieval architecture integrating kNN with MRF inference is then proposed. An experimental framework for the evaluation of the proposed architecture is illustrated, together with the discussion concerning the results that have been obtained.

2 Markov Random Fields

A *Markov Random Field* (MRF) is a Probabilistic Graphical Model (PGM) defined as the pair $\langle \mathcal{G}, \mathcal{P} \rangle$ where $\mathcal{G}$ is an undirected graph whose nodes represent random variables (we assume here discrete random variables) and edges represent dependency relations among connected variables. $\mathcal{P}$ is a probabilistic distribution over the variables represented in $\mathcal{G}$. In *Pairwise MRFs*: each edge $(X_i — X_j)$ is associated with a *potential* $\Phi_{i,j} : D(X_i) \times D(X_j) \to \mathbb{R}^+ \cup \{0\}$; here $D(X)$ is the domain (i.e. the set of states or values) of the variable X.

In a MRF, the distribution $\mathcal{P}$ factorizes over $\mathcal{G}$, i.e.:

$$\mathcal{P}(X_1 \dots X_n) = \frac{1}{Z} \prod_{i,j} \Phi_{i,j}(X_i, X_j)$$

where $Z = \sum_{X_1 \dots X_n} \prod_{i,j} \Phi_{i,j}(X_i, X_j)$ is a normalization constant called the *partition function*.

A special case of pairwise MRFs are *metric MRFs*, where all nodes take values in the same label space V, and a distance function $d : V \times V \rightarrow \mathbb{R}^+ \cup \{0\}$ is defined over V; the edge potentials take then a log-linear form

$$\Phi_{i,j}(x_i, x_j) = \exp(-w_{ij} d(x_i, x_j))$$

given that x_i is a value or state of variable X_i, and $w_{ij} > 0$ is a suitable weight stressing the importance of the distance function in determining the potential.

Concerning inference, we are interested in the computation of the posterior probability distribution of each single unobserved variable, given the evidence. In this paper, we will resort to *mean field inference*, a variational approach where the target distribution is approximated by a completely factorized distribution (Weiss 2000). This algorithm is implemented in the MATLAB UGM toolbox (Schmidt 2007) that we have exploited in our experimental analysis.

3 Characterizing Case Adaptability

In the standard CBR process, because of the similarity assumption we expect that the solutions of similar cases are also similar. If this assumption is not valid, kNN retrieval can result in a set of unuseful cases, since their adaptation level with respect to the query is unsuitable (i.e., no adaptation mechanism can be adopted with a reasonable effort). In order to circumvent this problem, we need to exploit adaptation knowledge during retrieval, by resorting to some abstract notion of adaptation space (see (Leake, Kinley, and Wilson 1997; Bergmann et al. 2016)). Let us assume, as done in (Stahl and Schmitt 2002; Abdel-Aziz, Strickert, and Hüllermeier 2014), that the solution space is equipped with a similarity or a distance metric.

Let $c_i(sol)$ be a boolean condition on a solution sol; we say that sol has an adaptation level i, if and only if $c_i(sol)$ is true. For the sake of simplicity, we assume in the following that each case has only one solution; we can then talk about the adaptation level of a case as the adaptation level of its solution.

Example 1. Suppose we have a case library containing the description of some used cars for sale (this corresponds to the case study described in section 6); let the solution of each case be the selling price of the car. Consider now a customer with a specific budget b and the possibility of a loan l; the following adaptation levels could be defined:

- Level 1: $c_1(price) \equiv (price \leq b)$
- Level 2: $c_2(price) \equiv (b \leq price \leq b + l)$
- Level 3: $c_3(price) \equiv price > b + l$

For instance, we can consider a solution with level 1 to be reusable, a solution with level 2 to be revisable, and a solution with level 3 to be non-adaptable. Solutions at levels 1 and 2 can then be considered as adaptable.

Given the above characterization, cases having similar solutions tend to have similar adaptation levels with respect to a given query. Adaptation levels reflect a different adaptation effort or cost, and can be tought as values of a linear feature

```
Input: #adapt_levels; st > 0
Output: a metric MRF
  MRF ← empty graph
  for each case c do
    add node c to MRF
  end for
  for each node n do
    num_states(n) ← #adapt_levels
  end for
  for each pair of nodes (n, m) do
    if sim_s(n, m) > st then
      add edge e = (n, m) to MRF
    end if
  end for
  for each edge e = (n, m) do
    s ← sim_s(n, m)
    for i = 1 ... #adapt_levels do
      for j = 1 ... #adapt_levels do
        Φ_{n,m}(i, j) ← exp(−s |i − j|)
      end for
    end for
  end for
```

Algorithm 1: MRF construction.

(Wilson and Martinez 1997). It is then straighforward to define their similarity as a measure of closeness of two levels on a linear scale (see section 4 for the details).

In the following, we propose to characterize the principle "similar solutions imply similar adaptation levels" by means of a metric MRF built on the case library; results from kNN retrieval are then used in order to provide evidence on the MRF, regarding the adaptation level of the retrieved cases. By computing the posterior probability of the adaptation levels of each case in the library, we can then suggest more potentially adaptable cases, and to finally reject cases that are likely to be unadaptable. Section 4 will discuss the details.

4 A Framework for Retrieval based on MRF Inference

Given a case library of stored cases with solutions, we first construct a metric MRF. Let `#adapt_levels` be the number of different adaptation levels, sim_s be the similarity metric defined over the solution space and st be a threshold of minimal similarity for solutions. Algorithm 1 shows the pseudo-code for the construction of the MRF.

Nodes of the MRF have the possible case adaptation levels as states (in the simplest situation they are binary nodes with states "0: adaptable" and "1: non-adaptable"), and they are connected only if the corresponding cases have a sufficiently large solution similarity. (greater than the threshold st). Concerning edge potentials $\Phi_{n,m}(i, j)$, since states represent adaptation levels that are linearly ordered, a simple distance metric as the absolute value of their difference is sufficient. When a given node assumes a specific state, connected nodes tend to assume close state values with a high probability (we expect cases having similar solution

Input: $al(1) \dots al(k); cond()$
Output: a set of (cases, adaptation levels) pairs (c, al)

```
for i = 1 : k do
  set_evidence(i, al(i))
end for
Bel[ ] = MRF_Inference
for each not retrieved case c do
  if cond(Bel[i]) then
    output (i, cond(Bel[i]))
  end if
end for
```

Algorithm 2: MRF inference for adaptable cases retrieval.

to have a very close adaptation level). Moreover, the more similar the solutions of the cases, the stronger this effect should be; this is the reason why we use solution similarity $s = sim\_s(n, m)$ as a weight for the metric potential.

Once we know the adaptation level of some of the stored cases, MRF inference can be used to propagate this information in the case library. The idea is to start from a standard kNN retrieval, followed by the usual reuse and revise steps. The results of the reuse/revise phases are used as input for MRF inference. Algorithm 2 details this process. Let $al(i)$ be the adaptation level of the i-th retrieved case (through kNN retrieval); for each retrieved case, its adaptation level is set as evidence in the corresponding node of the MRF. Inference is then performed and the posterior probability of each MRF node is computed into the multidimensional vector Bel. $Bel[i]$ is the posterior distribution or *node belief* of node i; $Bel[i]$ is a l-dimensional vector (where l is the number of adaptation levels) such that $Bel[i, j]$ is the probability of node i being in state j given the evidence. Input parameter $cond()$ is a function testing a condition on the node belief; if this condition is not satisfied, it returns `false`, otherwise it returns the state of the node (i.e., the adaptation level) for which the condition is satisfied. Algorithm 2 finally ouputs a set of cases with their adaptability level.

Example 2. Suppose we want to determine for each (non retrieved) case the most probable adaptation level, then we will set $cond(Bel[i]) = \arg\max_j Bel[i, j]$. In this case every case has a potential adaptation level and we can consider it for the next actions: for instance, we could be interested only in the most easily adaptable cases, and if 1 is the minimum adaptation level, we will select only those nodes i for which $cond(Bel[i]) = 1$

Consider now a more complex condition: suppose we consider as interesting any adaptation level from 1 to a, and suppose that we want to be pretty sure about the adaptability of the case. We could set a probability threshold pt and to require that

```
if Bel[i, 1] + ... Bel[i, a] > pt then
  cond(Bel[i]) ← a
else
  cond(Bel[i]) ← false
end if
```

In this case we are collapsing all the adaptability levels from 1 to a into a unique level (the choice of a is completely arbitrary here, and any other label would be fine as we no longer need to distinguish them); in case the required confidence on adaptability is not reached, we will simply ignore the case. Of course, several other implementations of the $cond()$ function can be devised.

5 An Integrated kNN/MRF Architecture

The problem of retrieving "useful" cases with respect to a given query is characterized by two different aspects: the structural similarity between the query and the retrieved case (addressed by kNN retrieval), and the adaptability to the query of the retrieved solution (addressed by MRF inference); this means that the cases of interest are those which are sufficiently similar to the query, while having an adaptable solution. We call them "positive" cases. We would like the retrieval to return only positive cases, possibly with a large structural similarity and with a low adaptability cost. While cases retrieved through kNN do not have the guarantee of being adaptable, cases retrieved through MRF inference are more likely to be adaptable, but they do not have any guarantee of being sufficiently similar (from the structural point of view) to the query. Moreover, the reason why retrieval is often restricted to a set of k cases, is because it is in general unfeasible to take into consideration all the positive cases: considering all the cases returned by MRF inference may lead to an unreasonably large number of cases to be managed.

In the following, we will consider a specific retrieval architecture combining kNN and MRF based retrieval trying to address the above mentioned issues, and we will evaluate it on a given case study. We start with kNN retrieval; if all the k retrieved cases are actually adaptable, then the process terminates with such k cases as a result. On the contrary, let $0 \leq k' < k$ be the number of adaptable cases retrieved by kNN; MRF inference is performed as shown in Algorithm 2, then the top $k - k'$ cases in descending order of structural similarity are returned, from the output of Algorithm 2.

The main idea underlying this architecture is that k is the desired output size. In case kNN retrieval tangles with the solution similarity problem, we complement the retrieval set with some cases that are likely to be adaptable. Since they are selected by considering their structural similarity with respect the query, we also maximize the probability of such cases being positive. In order to evaluate the effectiveness of such architecture, we set up an experimental framework described in section 6, and which results are reported in section 7.

6 Experimental Framework

We consider the "used cars" dataset related to the problem of suggesting suitable used cars to a requesting customer[1]. We have produced a sample of 350 cases having the following features: `Age`, `Miles`, `Doors`, `Power`, `Speed`, `CCM`,

[1]This dataset was delivered with an old version of the MyCBR tool (Bach and Althoff 2012) called CBRWorks.

`Price`. We consider the first six features to be the case description, and the feature `Price` to be the solution of the case[2]. Similarity measure for both case descriptions and solutions is computed as $s(i,j) = \frac{1}{1+d(i,j)}$ where $d()$ is the standardized Euclidean distance ($0 < s(i,j) \leq 1$).

In order to model an adaptability criterion, we consider the situation described in Example 1 of section 3: given a budget b and a possible loan l, we get three different adaptation levels depending on the retrieved price. Moreover, we adopt the evaluation criterion defined in Example 2 of section 4, by considering level 1 and 2 as adaptable and level 3 as non-adaptable[3]. We set the threshold $pt = 0.9$ (a case is adaptable if the probablity of being either reusable or revisable is greater than 90%), while a "positive" case is defined as a case having a structural similarity with the query greater or equal to a given threshold thr, and such that its adaptability level is either 1 or 2. We define the following measures for retrieval evaluation:

- True Positive (TP) cases: positive cases that are retrieved;
- True Negative (TN) cases: missed (non retrieved) cases that are not positive;
- False Positive (FP) cases: retrieved cases that are not positive;
- False Negative (FN) cases: positive cases that are missed.

We consider the usual notions of *precision* (p), *recall* (r), *accuracy* (a) and *F-score* (F).

$$p = \frac{TP}{TP+FP} \quad r = \frac{TP}{TP+FN}$$

$$a = \frac{TP+TN}{TP+TN+FP+FN} \quad F = 2\frac{prec\ rec}{prec+rec}$$

With respect to kNN, MRF based retrieval tends to move some cases from FN to TP, but with the risk of moving cases from TN to FP as well. What we expect is then an increase in the recall with a corresponding decrease in precision. To measure if this trade-off is well balanced we consider F: an increase of F means that increase in recall is worth the decrease in precision. If also accuracy increases, this further confirms that MRF based retrieval has been useful.

7 Experimental Results

We consider two different kind of queries: a query involving a small set of features, loosely correlated with the solution, and a query involving almost all the features of the cases. In the first situation, we expect the structural similarity not being a suitable proxy for solution similarity, increasing then the probability of retrieving non-adaptable cases via kNN. We adopted the thresholds $st = 0.9$ (see Algorithm 1) and

[2]All the features in this dataset are linear, however the generalization to the situation where also nominal features are present is straightforward.

[3]This criterion is just a constraint on the solution, and in principle it can be checked during kNN retrieval; it has been defined in order to set up a simple experimental framework. In real-world applications checking adaptability during retrieval could be impracticle or even impossible.

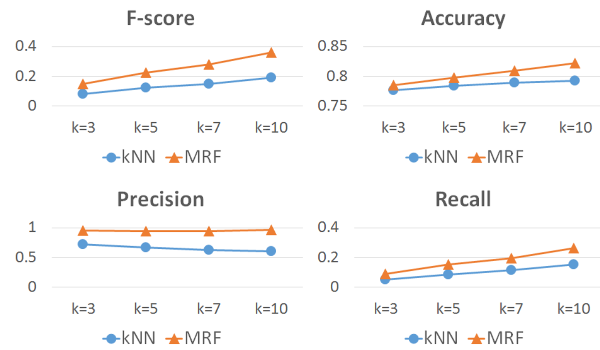

Figure 1: Small feature set, $thr = \mu_c$, $b + l = \mu_p$

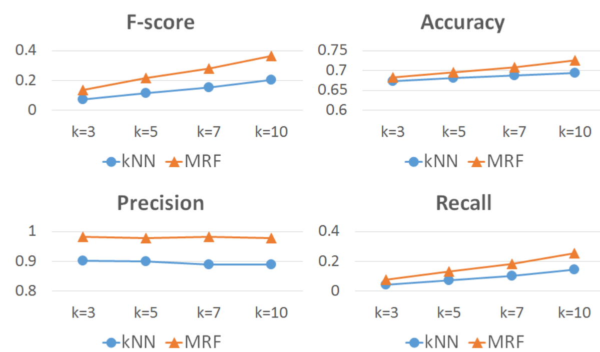

Figure 2: Small feature set, $thr = \mu_c$, $b + l = 2\mu_p$

$pt = 0.9$ (see Example 2 in section 4). A 10-fold cross validation has been performed in each experiment, in order to get 10 different test sets containing the queries, and 10 different corresponding case libraries. The construction of the metric MRF (Algorithm 1) and the MRF inference (Algorithm 2) have been implemented in MATLAB by exploiting the UGM toolbox (Schmidt 2007). In particular, we resort to mean field variational inference as previously mentioned[4]. We finally consider two other evaluation parameters: the average structural similarities among the cases in the case library μ_c, and the average value μ_p of the solutions (the average price of the cars in the case library). We have used these parameters in order to vary the structural similarity threshold and the adaptation criterion in the experiments.

The first set of experiments concerns queries restricted to features `Doors` (number of doors) and `CCM` (engine capacity). We have performed 4 runs of 10-fold cross validation, by varying parameters thr and $b+l$. We have measured F, a, p and r for different values of k, namely $k = 3, 5, 7, 10$. The average values of the performance measures are reported in figure 1 to figure 4 for different values of parameter k.

We can notice that MRF outperforms kNN in all experiments with regards to almost every measure (with the exception of p in figure 4). In figure 1, MRF is able to keep p close

[4]Comparable results have been obtained by using Loopy Belief Propagation (Weiss 2000).

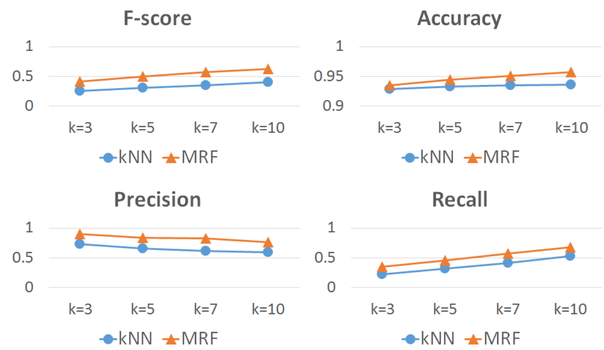

Figure 3: Small feature set, $thr = 2\mu_c$, $b + l = \mu_p$

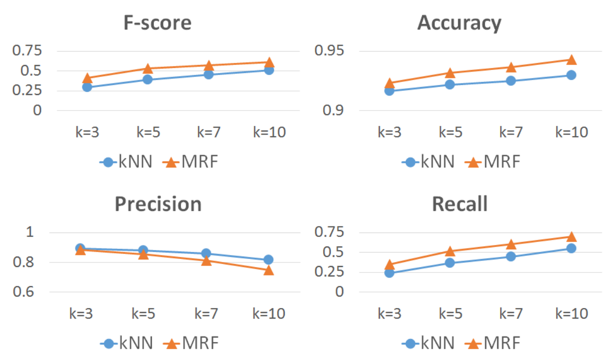

Figure 4: Small feature set, $thr = 2\mu_c$, $b + l = 2\mu_p$

to 1, but r is quite low, increasing as k increases. This is explained by the fact that there are a lot of cases sufficiently similar to the query and potentially adaptable; despite that, kNN alone has difficulty in retrieving positive cases (query features are not well correlated with the solution), and the integration with MRF inference is fruitful.

Similar situation is reported in figure 2, where we add more adaptable cases to the case library; there is a small improvement in p with a small decrease in a due to the presence of more FN cases. MRF shows a definite better balance of p and r than in kNN. Figure 3 concerns a situation where we restrict the set of positive cases by increasing the minimal required structural similarity; a and r become larger, since there are less FN and more TP. Precision shows a slight decrease as k increases, due to the fact that the presence of less positive cases provides an increase of FP cases. From the positive side, F is larger than in the previous settings. Figure 4 confirms that; however results concerning p deserve some attention. With respect to figure 3 there are more adaptable cases, so the drop of p as k increases is again explained by the presence of more FP cases; however, the value for kNN is larger than in the previous setting, because the presence of more adaptable cases limits the number of FP. MRF is worse than kNN in precision, since when it adds more adaptable cases to the retrieval set there is a larger probability of getting cases that are adaptable, but not sufficiently similar from the structural point of view. In conclusion, with

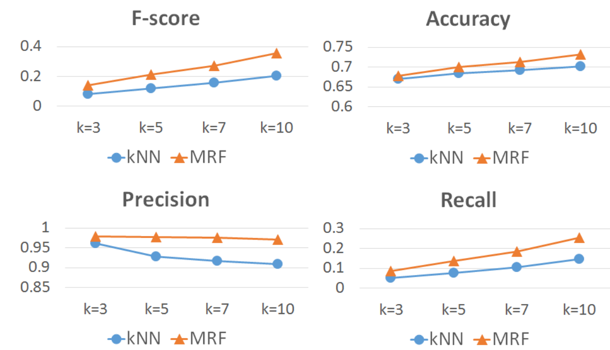

Figure 5: Augmented feature set, $thr = \mu_c$, $b + l = 2\mu_p$

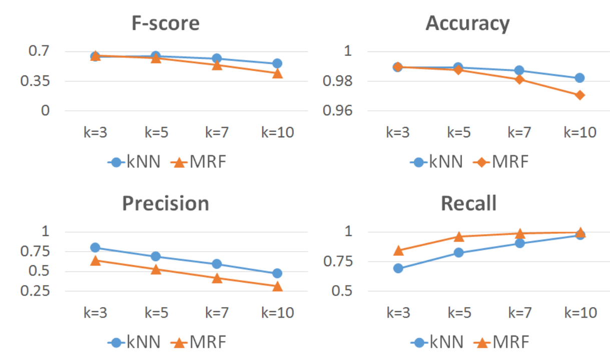

Figure 6: Augmented feature set, $thr = 2\mu_c$, $b + l = 2\mu_p$

a very restricted requirement for structural similarity and with a large number of adaptable cases, MRF inference can be penalized in terms of precision.

The second set of experiments concerns queries involving more features and in particular `Age` (number of years), `Miles` (miles traveled), `Power` (engine power) and `CCM` (engine capacity). All these features (both taken as alone and together) are quite correlated with the car's price, so we are modeling a situation which is potentially unfavourable for MRF. We report here 2 runs of 10-fold cross validation, by varying parameters thr, and by fixing $b + l = 2\mu_p$. The average results are reported in figure 5 and figure 6,

In the setting corresponding to figure 5 we still notice that MRF outperforms kNN, even in presence of features which are significantly indicative of the solution, with considerations that are similar to those obtained from figure 1 and figure 2. In contrast, results reported in figure 6 show a better performance of kNN, with the exception of r, where however both kNN and MRF tend to reach a perfect recall as k increases. On the other hand, the other performance measures show a decrease as k increases. Indeed, since there are less positive cases, more FP are introduced as k increases. These results show a situation where MRF introduces more FP than kNN alone. The decrease of F for both methods indicates that a large k is not suitable in this setting.

8 Conclusions

We have discussed a novel framework for enhancing kNN retrieval in CBR systems, by addressing the problem of retrieving cases that are likely to be easily adaptable. We have proposed to exploit a metric MRF model, taking into account the relationships among cases having similar solutions. We have introduced the notion of adaptation level, representing the effort a case solution requires in order to be adapted to the current query. We have evaluated a kNN/MRF integrated retrieval architecture where cases retrieved from standard kNN are complemented with cases retrieved through MRF inference. Experimental results, evaluated in terms of accuracy, precision and recall of the retrieval, suggest that the integrated architecture can provide advantages in several situations; in particular, when the query is underspecified and some features significantly correlated with the solution are missing, exploiting MRF inference can increase recall and accuracy without paying a big price in terms of precision. On the other hand, when the problem requires the retrieval of cases with very high structural similarity and a large number of adaptable cases is available, then MRF inference can be penalized in terms of precision; in this case, kNN alone can provide more effective results. We have reported the results of an experimentation concerning a dataset of 350 cases representing used cars, and an implementation of MRF inference based on mean field approximation. It is worth noting that, the approach is likely to scale-up, since a reasonable MRF model of the case library tends to have multiple connected components. This is due to the fact that only cases having close solutions are connected. Inference on such MRF components can be performed independently, and if they involve a limited number of nodes, even exact inference may be attempted.

Other researchers have already investigated the integration of CBR and PGMs for retrieval, usually by concentrating on directed models like Bayesian Networks (BN). In (Aamodt and Langseth 1998), a BN model is coupled with a semantic network to exploit statistically sound contribution to case indexing and retrieval. BN-based retrieval is triggered by the introduction of the observed features as evidence, and cases can be set in the `on` state and retrieved if the posterior probability of such state exceeds a given threshold. Recent advances in this setting are presented in (Nikpour, Aamodt, and Bach 2018) in the context of the BNCreek system which applies a Bayesian analysis aimed at increasing the accuracy of the similarity assessment. These approaches have some resemblance with our proposal, but they concentrate only on structural similarity and there is no attempt to address the problem of adaptation-guided retrieval. To the best of our knowledge, no approach integrating PGM and CBR directly tackle the problem of the validity of the similarity assumption.

References

Aamodt, A., and Langseth, H. 1998. Integrating Bayesian networks into knowledge-intensive CBR. In *AAAI Workshop on Case-Based Reasoning Integrations*, 1–6.

Abdel-Aziz, A.; Strickert, M.; and Hüllermeier, E. 2014. Learning solution similarity in preference-based cbr. In *Proc. ICCBR 2014, LNAI 8765*, 17–31. Springer.

Bach, K., and Althoff, K. 2012. Developing case-based reasoning applications using myCBR3. In *Proc. ICCBR-12, LNAI 7466*. Springer. 17–31.

Bergmann, R.; Muller, G.; Zeyen, C.; and Manderscheid, J. 2016. Retrieving adaptable cases in process-oriented Case-Based Reasoning. In *Proc. 29th FLAIRS 2016*, 419–424. AAAI Press.

Diaz-Agudo, B.; Gervas, P.; and Gonzales-Calero, P. 2003. Adaptation guided retrieval based on formal concept analysis. In *Proc. ICCBR 2003, LNAI 2689*. Springer. 131–145.

Leake, D.; Kinley, A.; and Wilson, D. 1997. Case-based similarity assessment: estimating adaptbility from experience. In *Proc. 14th AAAI 97*, 674–679. AAAI Press.

Murphy, K. P. 2013. *Machine Learning: a probabilistic prespective*. MIT Press. chapter Undirected Graphical Models (Markov Random Fields), 661–705.

Nikpour, H.; Aamodt, A.; and Bach, K. 2018. Bayesian-supported retrieval in BNCreek: A knowledge-intensive case-based reasoning system. In *Proc. ICCBR 2018, LNAI 11156*, 422–437. Springer.

Portinale, L.; Torasso, P.; and Magro, D. 1997. Selecting most adaptable diagnostic solutions through Pivoting-Based Retrieval. In *Proc. ICCBR97, LNAI 1266*. Springer Verlag. 393–402.

Schmidt, M. 2007. UGM: A Matlab toolbox for probabilistic undirected graphical models. http://www.cs.ubc.ca/~schmidtm/Software/UGM.html.

Smyth, B., and Keane, M. T. 1998. Adaptation-guided retrieval: questioning the similarity assumption in reasoning. *Artificial Intelligence* 102(2):249 – 293.

Stahl, A., and Schmitt, S. 2002. Optimizing retrieval in CBR by introducing solution similarity. In *Proc. IC-AI'02)*. CSREA Press.

Weiss, Y. 2000. Comparing the mean field belief propagation for approximate inference in MRF. In *Advanced Mean Field Methods: Theory and Practice*. MIT Press. 229 – 240.

Wilson, D., and Martinez, T. 1997. Improved heterogeneous distance functions. *Journal of Artificial Intelligence Research* 6:1–34.

Xiong, N., and Funk, P. 2006. Building similarity metrics reflecting utility in case-based reasoning. *Journal of Intelligent & Fuzzy Systems* 17(4):407–416.