The Case for Circularities in Case-Based Reasoning

Adwait P. Parsodkar[1], Deepak P[1,2], and Sutanu Chakraborti[1]

Indian Institute of Technology, Madras, Chennai, India
Queen's University Belfast, Belfast, UK
{cs20d404,sutanuc}@cse.iitm.ac.in deepaksp@acm.org

Abstract. In this paper, we demonstrate that diverse CBR research contexts share a common thread, in that their origin can be traced to the problem of *circularity.* An example is where the knowledge of property A requires us to know property B, but B, in turn, is not known unless A is determined. We examine the root cause of such circularities and present fundamental impossibility results in this context. We show how a systematic study of circularity can motivate the quest for novel CBR paradigms and lead to novel approaches that address circularities in traditional CBR retrieval, adaptation, and maintenance tasks. Furthermore, such an analysis can help in extending the solution of one problem to solve an apparently unrelated problem, once we discover the commonality they share deep down in terms of the circularities they address.

Keywords: Circularity · Truth Discovery · Knowledge Containers

1 Introduction

Case-Based Reasoning (CBR) is an inherently cyclical problem-solving paradigm where problem-solving and learning go hand in hand [1]. Given a target problem to solve, relevant past cases are retrieved, and their solutions are adapted to propose a solution. If the solution works, a new case comprising the target problem and the verified solution is stored in the case base. With each new case being added, the system gets more competent, and this constitutes learning. If the system is good at problem-solving, the cases fetched using retrieval and the solutions proposed using adaptation stand a better chance of being useful – such successful episodes of problem solving help in growing the case base and making it more competent. However, it is also that true problem solving cannot be effective unless the case base is competent. Such circularities abound more generally in the broader context of knowledge representation and reasoning: the effectiveness of the reasoning mechanism critically depends on the knowledge representation, but then, the knowledge itself is acquired by way of reasoning.

The circular dependence of problem solving and learning in CBR is just an instance of circularities that manifest in diverse research sub-themes of CBR. This paper is intended to be an exploratory one, where we attempt to examine the genesis, the nature, and implications of such circularities. In particular, we attempt to substantiate our central claim: whenever we run into a circularity, it makes sense not to sweep it under the carpet. A critical examination can help us gain foundational insights on the nature of a problem, and spawn interesting research themes that can mitigate, if not resolve, the underlying circularity. At the very least, we emerge more well-informed of the unwritten assumptions behind existing approaches.

The structure of the paper is as follows. Section 2 presents the big picture, where we position the circularities observed in CBR in the context of a wider class of circularities that arise in the context of knowledge representation, and demonstrate how they share common origins. We introduce the notion of "representation gap", a term we use to denote the disparity between the knowledge as represented in a CBR system, and the actual experiences they attempt to encode. Our central thesis is that representation gap is the single most fundamental source of circularities. In Section 3, we show how an in-depth appreciation of the knowledge representation gap can lead to questioning the fundamental assumptions of traditional CBR, which are often taken to be granted, and thereby motivate research into non-conventional paradigms in CBR. Section 4 is about the more commonplace ways of dealing with circularity, where we operate within the framework of traditional CBR and address circularities as and where they surface. One auxiliary goal of our paper is to pave the way for a unifying framework, where apparently unrelated research problems in diverse areas of CBR are found to share commonalities in terms of the underlying circularities that they set out to resolve. Such revelations can open up opportunities for transplanting or adapting the solution of one problem to another. We present some examples to illustrate this in Section 5. Finally, Section 6 concludes the work and reflects on potential research directions that it can open up.

2 Circularities: The Big Picture

An interesting case of circularities is one of loops and self-reference in dictionary definitions [21]. For example, the definition of the word "green" makes use of the word "leaf" (to give examples of objects that are green), but "leaf", in turn, is defined using the word "green". Such circularities arise because of limits of language: the perception of the colour "green" cannot be communicated to a visually challenged person using language alone. This is an instance of a wider class of circularities that originate because of the "representation gap": the fact that the perceptual grounding of symbols [14] is not contained in the language itself. In the context of CBR, we can show that irrespective of the richness of knowledge contained in vocabulary, case base, similarity, and adaptation [32], the representation is, at best, a surrogate for the experiences that give birth to the cases and the underlying design choices are in the mind of the designer, not in the system itself. This places a natural upper bound on the effectiveness of the system when viewed as an autonomous agent. In Sections 3 and 4, we deliberate on ways of mitigating the effects of circularities whose genesis can be traced back to the representation gap.

Let us examine more closely the nature of the representation gap, its theoretical implications, and how it manifests as circularities in the context of CBR. Formal arguments that demonstrate the limits of language were advanced by the noted philosopher of language Ludwig Wittgenstein. In Wittgenstein's words [39]:

> *"In order that you should have a language which can express or say everything that can be said, this language must have certain properties; and when this is the case, that it has them can no longer be said in that language or any language."*

While this is cryptic on the surface, the essence can be distilled using a simple example. The coupling of the word "green" to the perceptual experience of the colour "green" in language L_1 (say English) is not contained in L_1 – so one would require a second language L_2 to establish that association. But any symbol that constitutes the alphabet of L_2 suffers from the same limitation, in that it requires a third language L_3 to establish the grounding of its symbols. L_3 similarly calls for yet another language L_4, and so on, giving rise to an infinite regress. This is an intuitive sketch of the argument used to show that no language is self-contained.

Closely related to the unsurmountable limits of language are problems originating from an (unwritten) assumption in science that an observer can study nature as an objective reality outside him. When the mind studies itself, however, the observer becomes identical to the observed, leading to a fundamental circularity. For an engaging account of the implications of such circularities see [2]. Even in physics which studies matter and shies itself away from studying mind, and where the traditional view has been that the observer is distinct from the observed, the advent of quantum mechanics has dealt a heavy blow, questioning its foundations. In particular, we have impossibility results like the Heisenberg's uncertainty principle. Interestingly, Heisenberg is categorical in attributing aspects of his results to fundamental limits of languages as revealed by Wittgenstein [39]. In the words of Heisenberg:

> *"Words have no well-defined meaning. We can sometimes by axioms give a precise meaning to words, but still we never know how these precise words correspond to reality, whether they fit reality or not. We cannot help the fundamental situation - that words are meant as a connection between reality and ourselves - but we can never know how well these words or concepts fit reality. This can be seen in Wittgenstein's later work."*

Interestingly, a recent paper by Bender and Koller [4] has argued that despite the impressive achievements in the field of Large Language Models (LLMs), since the latter only use form as training data, it can never lead to learning of meaning, where meaning is the relation between linguistic form and communicative intent. This is in strong agreement with our observations on the representation gap.

The implications in the context of CBR follow naturally. The knowledge of a CBR system can never be self-contained. Irrespective of how rich its knowledge containers [32], i.e., the case base, vocabulary, similarity, and adaptation are, a CBR system will always fall short of what we would ideally expect of it. The case base and the vocabulary used to express and index the same can never capture the totality of experiences of problem-solving in any domain, nor can the similarity and adaptation knowledge capture all the diverse settings in which such experiences can and cannot be applied.

This representation gap manifests itself in various kinds of circularities in the context of CBR. One potent example is related to the inherent challenge in defining the utility of a case with respect to a target problem [5]. In particular, the statement of circularity is *Given a target problem, a case should be retrieved only if it is useful in solving the target problem. However, we cannot ascertain the utility of a case with respect to the target problem unless it is retrieved in the first place.*

Since it is impossible to resolve this circularity, we often bypass it by using similarity as a surrogate for utility. Unlike utility which is an "a posteriori" criterion in that we can know the utility only after we have solved the target problem, similarity is "a priori" in nature: we can estimate similarity and use it as the basis for retrieving cases while acknowledging that the cases thus retrieved are not necessarily the ones most useful in the "a posteriori" sense (for details, see Section 4.1 in [5]).

A second example is in the context of purely introspective learning, where similarity knowledge is acquired (inductively) in a purely bottom-up fashion from the cases, and there is no external knowledge separate from the case repository that provides that knowledge. Let us consider the case where cases are represented in terms of attribute values. We encounter a circularity of the kind - *Two cases are regarded as similar when they have similar attribute values. However, two attribute values are regarded as similar when they occur in similar cases.* It is interesting to note that there is no such circularity in settings that are not purely introspective, such as those in which domain experts explicitly provide local similarity functions.

As would be expected, such circularities are commonly encountered in several contexts in machine learning, which is largely based on inducing hypotheses introspectively from data. For example, in learning distributional measures of word similarity (or more appropriately, word relatedness) from a given document corpus, we exploit underlying word co-occurrence patterns. In particular, we observe: *Two words are similar if they occur in similar documents*[1]*. However, two documents are similar if they have similar words.* The circularity inherent here is isomorphic to the circularity we observed in the context of cases and attribute values. In text mining literature, several approaches have been proposed for estimating the relatedness of words and documents by addressing such circularities. Examples are : (a) factor-analytic approaches like Latent Seman-

[1] Instead of documents we can choose to consider paragraphs, or some other unit of discourse.

tic Analysis (LSA) [8] (b) the use of Expectation Maximization for parameter estimation as in Probabilistic LSA, (c) distributional representations which explicitly model homophily [25], and (d) the SimRank algorithm [19], which can be viewed as an extension of the PageRank algorithm [29]. Interestingly, we find that the use of such approaches for addressing the problem of circularity is not surprising: PageRank, which is designed to identify web page importance based on the hyperlink structure of the web, for instance, already solves a circularity: *A web-page is important if it is pointed to by several important web-pages.* Similarly, the E and M steps in the EM algorithm are known to handle circularities inherent in parameter estimation from incomplete data. In particular, we can only estimate parameters when the data is completed; but the data cannot be completed unless we know the parameters [24]. The discussion above points to the fact that if we handle circularity effectively in one context, the solution can be reused in addressing circularities in other contexts as well. We will have more to comment on this, especially in the context of CBR, in Section 5. When knowledge is acquired externally, circularities of the kind shown above do not appear. For example, if relatedness between two words is externally provided by a source like WordNet [26], word similarities are no longer circularly dependent of document similarities. However, even while using WordNet, there would still be issues in grounding meanings of words. In the context of Word Sense Disambiguation, which is used to identify which of the several senses of a polysemous word (like "bank") must be chosen given some context words, we encounter a fresh kind of circularity: *the correct sense of a word W_1 can only be known if we know the correct sense of a context word W_2, but the disambiguation of W_2 needs W_1 to be disambiguated in the first place.* Thus the circularity takes a different form, but does not completely disappear. This is not surprising, given our central argument that the representation gap can, at best, be minimized; it can never be done away with.

A concept very closely related to the representation gap is that of "structure function correspondence" [35]. In structural CBR (as with relational databases), the structure of a case is strongly representative of the functions that it can carry out, so the structure-function correspondence is strong. On the other hand, we have relatively weak structure function correspondence in textual CBR – this is because words in text interact with each other non-linearly to give rise to emergent "meaning" that may not be carried by any of the words in isolation. Poetry, satire or humour are at the extreme end of the spectrum, where structure function correspondence is very weak. The weaker the structure function correspondence, the larger the representation gap.

There is yet another perspective that is relevant: one that posits that circularity originates from the dichotomy of the whole versus parts. St. Augustine theorized that children learn the meanings of words by the following simple process: an adult points his finger to an object and says "chair", the child associates her visual perception of that object with the word "chair". This theory has glaring loopholes, however. Based on such a limited interaction with the adult, it is virtually impossible for the child to tell whether "chair" refers to "anything made of wood", "any object on which we can sit", or "any object with a flat surface", "the particular object being pointed to alone", or any of other such possibilities. Thanks to Ferdinand de Saussure and his theory of structuralism [10], we now know that this is not how word meanings are acquired. Words refer to concepts are abstract entities that can only be defined by means of their relationship to other concepts. For example, the word "chair" does not refer to any specific instance of a chair or to the collection of all chairs in the universe – rather, it is an abstract concept that can only be defined in terms of associations with other concepts that define its function ("sit"), constituents ("wood" or "steel"), the super-category or hypernym ("furniture"), subcategories ("easy chair" or "office chair"), and concepts that have similar functions ("stool" or "bench"), for example (refer [3] for a relevant discussion on the Protos system in the context of CBR). *An entity A cannot be ontologically defined without reference to a related entity B; the definition of B, in turn, requires A to be defined.* This reminds us of the mutual dependence of the words "leaf" and "green" in dictionary definitions and hence the circularity originating from structuralist assumptions, at its core, is closely related to circularity that stems from the representation gap.

3 Taking Circularity head on: Alternate Paradigms

The first way to address circularities emanating from the representation gap is to come up with novel paradigms that challenge the foundational assumptions behind traditional CBR systems. In this section, we present two illustrations of such paradigm-level changes.

The first is the existing holographic CBR paradigm [11,37]. The Holographic CBR paradigm was first proposed by Devi et al. [11] with the goal of making CBR more cognitively appealing. It was subsequently extended by Renganathan et al. [37]. In traditional CBR systems, cases can be added or deleted without affecting the rest of the case base. Human memories, in contrast, are far from passive – we tend to actively reorganize our experiences. To cite an example by Roger Schank [34], whose early work on the role of memory structures played an important role in shaping up the field of CBR, we tend to forget details of individual visits to restaurants (unless something very different from the norm happens) – instead the specifics gets "mushed up" to create more general structures that are effective in raising expectations about future trips, and seeking explanations in case the event of expectation failures. The departures from conventional models are not just at the level of high-level cognition; there are interesting findings from neuroscience as well. In particular, we now know that when portions of the brain are damaged or surgically removed, it is not that memories are wiped out completely. Instead, what we have is a hazy reconstruction of what was stored previously. This can be attributed to non-localized or holographic memories, where each component carries an imprint of the whole.

Holographic CBR draws inspiration from these ideas. We have seen in the last section that the part-whole dichotomy leads naturally to the representation gap which manifests itself in the form of circularities. Holographic CBR attempts to address the sharp divide between the whole and parts by making sure that each case is no longer a passive record of problem solving, but they house similarity and adaptation knowledge locally, proactively interact with each other in the course of problem solving in a distributed way, and offer resilience in the face of case deletion and the ability to generalize when cases come in. This is in sharp contrast to traditional CBR system, where cases are passive and the reasoner, which has access to similarity and adaptation knowledge, has centralized control of how cases are used for problem solving. A conventional CBR system is analogous to an organization where the manager runs the show himself, and each of his subordinates passively execute the job given to them, while being agnostic to the overall objective, or the competencies of their co-workers. In contrast, a holographic CBR system is analogous to an organization where the employees are aware of the overall goals and competencies of their colleagues, actively collaborate with each other in problem solving, possess the ability to take over responsibilities from their co-workers in case any of them leave the organization (case deletion) and can also hire new workers if needed (case addition). In the context of CBR, the central insight from this analogy, the closer the case representations are, to the intent for which they have been created in the first place, the lower the representation gap.

A second example of a paradigm level shift can potentially be made by use of quantum models. In the context of quantum mechanics, Heisenberg's uncertainty principle declares fundamental limits to the precision with which certain pairs of physical properties of a particle, such as position and momentum, can be known. Interestingly, Heisenberg attributes aspects of his results to Wittgenstein's work on limits on languages [31], which directly relates to the representation gap and consequently the problem of circularity as we have seen earlier. The mathematical framework of quantum physics has been exploited by Busemeyer [6] in the context of quantum models of cognition. The seminal work of Rijsbergen [33] in the area of quantum information retrieval is also a case in point. In future, we may see extensions of the quantum models pave way for novel paradigms that can effectively contain representation gap in CBR.

4 Addressing Circularity as and where it appears

Our discussion in the last section illustrated how a study of circularities at a fundamental level can motivate research into non-conventional paradigms in CBR. It should be noted, however, that occurrences of circularities are also observed to manifest within the current framework of CBR. In this section, we show how a few of these have been addressed in the literature.

4.1 Addressing the problem of modelling case utility

The issue of defining the utility of a case with respect to a target problem has been highlighted in Section 2 with emphasis on the fact that the similarity is only a proxy to its utility. The CBR literature recognizes the limitations of the use of similarity as a surrogate for utility and has proposed measures such as those based on diversity, adaptability, and compromise to complement similarity.

A diversity-conscious retrieval will allow the recommendation (of, say, houses) of heterogeneous characteristics that are relevant to the query. Such a strategy allows the system to learn about the user's preferences in situations of uncertainty (for instance, whether or not the user wants a house with an open layout). Compromise-driven retrieval, yet another approach to expose the users to a diversity of products, operates by providing the user with recommendations that make compromises across different attributes (price of the house, for instance). Rounds of interactions with a system employing a compromise-driven retrieval can reveal the attributes on which the user is willing to make compromises and allow navigation to products with, say, the minimum number of compromises.

An adaptation-guided retrieval takes into account the adaptability of cases to suit the query and retrieves cases that are not only close to the query but are also highly adaptable. As an illustrative example, while a laptop may not have on-paper specifications that closely resemble the user's requirements, if it can be customized to meet the user's needs precisely, it should ideally be retrieved.

It is critical to note that these measures, much like similarity, are "a priori" in nature thereby fundamentally limiting them from being good estimates of utility that is "a posteriori" in nature. This aligns with the impossibility result in Section 2 where we argue as to why the representation gap can never be bridged completely. However, a variety of "a priori" measures can be used to arrive at a better estimate of the utility of a case to a query, and hence serves as means to narrow the representation gap.

4.2 Reducing Representation Gap with better interfaces

An attribute-value style description of cases in CBR is fundamentally bound by the amount of details that can be captured about the corresponding entity. The same applies in the context of Information Retrieval where the query is not adequately expressive of the underlying intent of the user, and similarly, the document indices also fail to capture all the different ways in which it may potentially be found useful.

Attempts have been made in the CBR community to reduce this representational gap by exploiting user feedback to compensate for the lack of richness in the utility function. The provision of allowing user feedback of varying characteristics over the set of candidates can allow the system to learn the user's implicit preferences. In particular, feedback that highlights a user's preference for one product over another can provide valuable insights into the features favored by the user. This type of feedback can also be memoized to identify broader product preference trends among users with similar preferences [38]. The user may also *critique* a recommendation (a camera for instance), by requesting a product with better characteristics (higher resolution or lower price, for instance) [7]. Furthermore, a *more like this* [12] feedback mechanism allows the system to recommend options that are still similar to the selected product in aspects previously expressed by the user while offering a diverse range of options across other features. For instance, a user expressing a *more like this* feedback for a budget laptop can be further exposed to options from diverse brands in order to understand her brand preference.

In the light of the discussions in Section 2, these approaches improve structure-function correspondence and hence reduce the representation gap.

4.3 Narrowing Representation Gap by a way of novel retrieval formalisms

Case Retrieval Networks [20] is a spreading activation model that facilitates flexible and efficient retrieval in CBR. The network involves a set of information entities (the set of all attribute values), denoted by IE, and a set of case nodes. In order to reflect the non-orthogonality of the IEs, weighted connections, referred to as the *similarity arcs*, between IEs are introduced. The association between each pair comprising of an IE and a case is expressed via *relevance arcs*. Upon encountering a query that is viewed as a subset of all the IEs, activation is first spread within the IEs via the similarity arc. This is followed by the propagation of values to the cases by means of the relevance arcs. The cases are then sorted based on the aggregated values.

It is to be noted that the retrieval in CRN is dependent only on the input query (feature values, in the CBR context) and is based on (a rather optimistic) assumption that all the information in a case is captured by the totality of the information contained in the IEs. While this might be the case in situations with strong structure-function correspondence (such as a task involving the prediction of house prices), it often remains inapplicable in those with poor structure-function correspondence which is often characterized by a large representational gap. A book recommender system, one of the several settings where attribute-value representations fall short, is associated with the representation strategy exhibiting poor structure-function correspondence in that the attributes of a book are not necessarily indicative of its contents.

Collaborative filtering style approach [18] serves as a viable option to handle this issue that estimates relatedness between books, in our example, based on the relatedness of the users who rate them. This allows the incorporation of user preference knowledge into the system that is manifested in the CRN architecture via arcs between cases [35]. Notice that the integration of arcs between cases facilitates retrieval that is not solely committed to the features but takes into account knowledge from a host of users, thereby allowing the narrowing of the representation gap.

4.4 Case Reliability

In large-scale CBR applications, the knowledge contained in the case base in the form of cases can be sourced from an array of external sources such as datasets, websites, etc. Since these sources might cater to different quality standards, the resulting case base need not necessarily be composed of homogeneous quality cases. A case exhibiting poor quality in a classification setup might correspond to what is often referred to as noisy, in machine learning literature (e.g., noisy labels [27]). The presence of such cases in the case base can significantly impact the prediction performance of the system. While it is theoretically possible for a domain expert to manually assess these cases for quality, it is apparently infeasible in practice. Arriving at approaches that arrive at estimates of the quality of the cases, in the absence of labeled data, appears promising with regards to improving the efficacy of the reasoner.

Although measures have been proposed to identify cases of substandard quality ([23], for instance), they often rely on the underlying definition that a case is unreliable if it disagrees with its neighbors. Such approaches implicitly assign a homogeneous quality to all the neighbors - something that might rarely be true in several settings with unvalidated cases. We would, however, like to take into account the contribution of neighbors of a case in proportion to their quality while estimating the quality of a case. This leads to an immediate circularity - inferring the quality of a case requires the knowledge of the quality of the individuals in its neighborhood. RelCBR [30] proposes the following definition for the *reliability* of cases based on the following circular definition.

Definition 1. A case is reliable if it can be solved by its reliable neighbors.

The circularity resolution approach involves translating this statement into an objective function that is optimized for reliability scores using the projected gradient descent approach [30]. Empirical results demonstrate that this definition facilitates appropriate assignment of reliability to cases even if they appear in a neighborhood of noisy cases, a situation where the baseline methods arrive at misleading conclusions.

4.5 Case Acquisition in Textual CBR

In this section, we highlight a previously established approach, called Correlation and Cohesion Driven Segmentation (CCS) [28], which attempts to segment a set of textual incident reports into a case base. Specifically, each report is to be split into two components, namely, a problem component and a solution component. The basis for this approach is founded upon a circular theme expressed in the following statement.

Definition 2. Good segmentation of each of the documents can serve as a valuable means to understand the behavior of the problem and solution segments. Well-learned solution and problem segment behavior can facilitate the derivation of good segmentation.

Here, the *behavior* of the problem and solution segment is modeled using language and translation models, details of which can be found in [28].

This circularity, much like most others, appears like the “chicken egg” problem [9]. In the situation at hand, if one were to arrive at a good understanding of how the problem and solution components corresponding to documents look like, a good segmentation strategy would be required; however, a good segmentation strategy can only be arrived at when the system has a good understanding of the problem and solution components. The EM algorithm is employed to resolve this circularity. In that, the E-step estimates the posterior probabilities associated with each potential segmentation, given the language and translation models, while the M-step reconstructs these models based on the probabilities.

4.6 Enriching case descriptions

Traditionally, the similarity function in CBR often only takes into account the query input to the system alongside the problem parts of the cases, disregarding the useful information gathered by the reasoner over interactions with its users. The system should ideally minimize the number of interactions with the customer and should therefore place *better* products at the top of the recommendation list. Notably, customers tend to favor products with positive collective reviews, such as bestsellers, over products with favorable on-paper features alone.

Since the case representation often falls short of such information, enriching the representation strategy based on knowledge drawn from user interactions holds promise for better recommendations [38]. In particular, we want criterion (such as *extent of violence* in a movie domain) to reflect as attributes in the case representation. Given the information about the preference of a product over another based on multiple users in the context of the criterion, useful insights can be drawn about the involved products. It is important to note that merely tallying the number of times a product has been preferred over other products is not a reliable measure of its overall *preferability.* This is attributed to the fact that a preference over a highly sought-after product is worth much more than that over a moderately preferred one. The following definition can be used to quantify the preferability of a product given a context.

Definition 3. A case is preferable in a given context if it is preferred over other preferable cases by multiple users in the same context.

This circularity resembles the PageRank style of rank-ordering pages. The knowledge of the preferability of the products, the analog of the importance scores in PageRank, can now be incorporated in the case representation by treating the preferability score as a value to the criterion considered as a feature.

5 Transplanting solutions of one kind of Circularity to solve another

Following on from Section 4 that presents a range of circularities in CBR and approaches employed to resolve them, we now highlight approaches that have conventionally been used in diverse CBR tasks like retrieval, adaptation, and maintenance and have been agnostic to the circularities latent in the problem definitions. Highlighting the underlying statements of circularity can open up fresh research directions towards improving the efficacy of the reasoner. By taking into account the underlying circularity, we can arrive at improved variants of these approaches. Based on the discussion presented in this section, it may be noted that circularities associated with seemingly unrelated problems may have analogous resolution strategies.

5.1 Ensemble of Knowledge Containers

Consider a setting that requires assigning reliability to the reviewers of a conference, wherein each reviewer is tasked with rating the work they review over a maximum score of, say, 5. In the absence of a higher authority that assesses the reviewer, reviewers whose rating is close to the inferred *trustworthy score* of, say, 4 for a paper are deemed *reliable*. However, the score 4 is determined based on the reliability of the reviewers in the first place, leading to a circular dependence. Truth Discovery [22] deals with the task of assigning reliability scores to a set of potentially conflicting sources (reviewers, in our example) that provide values (scores out of 5) to objects (research papers) in the absence of any form of supervision. The fundamental principle of circularity is captured in the statement: *A source is reliable if it provides trustworthy solutions. A solution is trustworthy if it is supported by reliable sources.* We now highlight how ideas inspired by the Truth Discovery literature can be leveraged in the context of CBR, where an ensemble of knowledge containers of the same kind is available.

5.1.1 Ensemble of Similarity Functions

Literature in CBR has proposed multiple ways of arriving at a similarity function in situations where top-down expert knowledge is not accessible (see [15] for instance). However, no single method is consistently shown to outperform others across multiple datasets, so the varied knowledge captured in a variety of similarity functions can be leveraged to result in improved performance. A rudimentary approach for aggregating these can involve taking a weighted combination of the different similarity values. This, however, requires similar distribution to be associated with the similarity functions in order to prevent a similarity function that is liberal in the assignment of similarity scores from overpowering another that usually assigns low similarity scores.

It is, therefore, advantageous to aggregate the rankings generated by these similarity functions instead, in addition to assigning them reliability scores that are learned in a bottom-up fashion. Notably, the statement of circularity in Truth Discovery, is analogous to the situation at hand, wherein a ranker (similarity function) corresponds to a knowledge source that generates rankings of cases as a solution to an input query. The following definition can be used to arrive at the reliability of the similarity functions.

Definition 4. A ranker is reliable if it produces trustworthy rankings for several queries. A ranking is trustworthy if it is supported by reliable rankers.

5.1.2 Ensemble of Adaptation Strategies

A bottom-up approach for learning adaptation rules, particularly in numerical prediction tasks, is to use the *case difference heuristic* [13]. In particular, for a pair of neighboring cases (c_i, c_j), an adaptation rule can be generated that states *'A difference of $p_i - p_j$ on the problem side results in a solution side difference of $s_i - s_j$'* where p_i and s_i indicate the problem and solution part of the case c_i.

Apparently, methodologies that use this strategy as a basis for acquiring adaptation knowledge are susceptible to two primary setbacks. The first of these pertains to the possibility of a drastic increase in the number of adaptation rules. Further, while existing approaches tend to assign uniform reliability to each adaptation rule in the ensemble [17], we envisage that several adaptation rules learned from the data might be incorrect owing to their derivation from poor-quality cases. Thus, assigning reliability scores to these rules hold potential for improved efficacy of the system.

The method outlined in [16] involves quantifying the reliability of an adaptation strategy by evaluating the goodness of the modified solution it generates for several queries, possibly in a leave-one-out style strategy. Although this methodology yields good reliability estimates when the cases in the case base are accurately labeled, it could lead to erroneous conclusions when dealing with noisy cases. In particular, a rule may be deemed unreliable due to its inability to accommodate poor-quality cases. Consequently, a more appropriate approach would be to verify if other reliable adaptation strategies concur with the rule under evaluation, in which case, the rule can be rightfully considered reliable. Much like our discussion in Section 5.1.1, we foresee the applicability of the Truth Discovery style of circularity to gauge the reliability of the adaptation rules using the following definition.

Definition 5. An adaptation strategy is reliable if it produces trustworthy solutions for several queries. The solution corresponding to each of these queries is trustworthy if it is agreed upon by several reliable adaptation strategies.

5.2 Robust Estimates of Case Competence

In the context of lazy learners such as Case-Based Reasoners, the query response time is typically found to increase with an increase in the number of instances (cases in CBR) [36]. The underlying reason being the increase in time required for the computation of similarity (or equivalently, distance) between the query and (potentially) every case in the case base. The Footprint Algorithm [36] returns a subset of the entire case base, referred to as the footprint set, while preserving the problem-solving ability, but eliminating redundant cases, thereby improving the efficiency of the system.

The set of cases that constitute the footprint set is critically determined by scores assigned to them that indicate their problem-solving abilities. Notice that approaches that indicate the goodness of the case (reliability, for instance) are not indicative of the redundant nature of the cases. Instead, the Relative Coverage [36] measure is often employed to assess the competence of a case, assigning it a high value if it solves cases that are not solved by many cases.

Consider a sample example where a case c_1 solves a case c_2 that is also solved by c_3. The Relative Coverage of c_1 is only dependent on the number of cases it solves (such as c_2) and the count of cases with which its coverage sets intersect (like with c_3). Apparently, the Relative Coverage score of c_1 is independent of that of c_3. However, the competence of c_1 should ideally be high if c_3 has a low competence (heuristically speaking, is likely to be dropped, thus making c_1 likely the only one that may be able to solve c_2 in a compacted case base such as the footprint). In effect, the competence of c_1 can be known only when that of c_3 is known, while the determination of competence of c_3 requires the knowledge of competence of c_1, leading to a circular dependence. To capture this, we propose a measure of competence of a case grounded upon the following definition.

Definition 6. A case is highly competent if it solves several cases that are solved by cases of low competence.

6 Conclusion

In this paper, we outlined how circularities manifest in diverse contexts in CBR, and how such circularities are organically grounded within understandings of knowledge representations and insights from cognitive and neurological sciences. We illustrated how a systematic study of such circularities can help us in three distinct ways. Firstly, it can spawn research on novel paradigms. Secondly, when we question the assumptions behind approaches in traditional CBR that are used for specific tasks, such as those involved in retrieval, adaptation, and maintenance, we often find that there are underlying circularities, which, when resolved effectively, can potentially lead us to improved variants of these algorithms. Thus a study of this kind has the potential to spawn fresh research across a wide range of themes within CBR. Finally, we have also seen how such a unified perspective helps in reusing solutions to circularity in one setting to other areas that, on the surface, appear unrelated.

References

1. Aamodt, A., Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI communications **7**(1), 39—59 (1994)
2. Aharoni, R.: Circularity. WORLD SCIENTIFIC (Sep 2015). https://doi.org/10.1142/9805
3. Bareiss, E.R., Porter, B.W., Wier, C.C.: Protos: an exemplar-based learning apprentice. International Journal of Man-Machine Studies **29**(5), 549–561 (1988). https://doi.org/https://doi.org/10.1016/S0020-7373(88)80012-9
4. Bender, E.M., Koller, A.: "Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data". In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. pp. 5185–5198. Association for Computational Linguistics, Online (Jul 2020). https://doi.org/10.18653/v1/2020.acl-main.463
5. Bergmann, R. (ed.): Experience Management. Springer Berlin Heidelberg (2002). https://doi.org/10.1007/3-540-45759-3

6. Busemeyer, J.R., Bruza, P.D.: Quantum Models of Cognition and Decision. Cambridge University Press (Jul 2012). https://doi.org/10.1017/cbo9780511997716
7. Chen, L., Pu, P.: Critiquing-based recommenders: survey and emerging trends. User Modeling and User-Adapted Interaction **22**(1-2), 125–150 (Oct 2011). https://doi.org/10.1007/s11257-011-9108-6
8. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R.: Indexing by latent semantic analysis. Journal of the American Society for Information Science **41**(6), 391–407 (Sep 1990). https://doi.org/10.1002/(sici)1097-4571(199009)41:6<391::aid-asi1>3.0.co;2-9
9. Domingos, P.: The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World. Basic Books, New York (2015)
10. Elffers, E.: Saussurean structuralism and cognitive linguistics. Histoire Épistémologie Langage **34**(1), 19–40 (2012). https://doi.org/10.3406/hel.2012.3235
11. Ganesan, D., Chakraborti, S.: "Holographic Case-Based Reasoning". In: Watson, I., Weber, R. (eds.) Case-Based Reasoning Research and Development. pp. 144–159. Springer International Publishing, Cham (2020)
12. Ginty, L.M., Smyth, B.: "Comparison-Based Recommendation". In: Craw, S., Preece, A. (eds.) Advances in Case-Based Reasoning. pp. 575–589. Springer Berlin Heidelberg, Berlin, Heidelberg (2002)
13. Hanney, K., Keane, M.T.: Learning adaptation rules from a case-base. In: Lecture Notes in Computer Science. pp. 179–192. Springer Berlin Heidelberg (1996). https://doi.org/10.1007/bfb0020610
14. Harnad, S.: Categorical Perception. In: Nadel, L. (ed.) Encyclopedia of Cognitive Science, pp. 67–4. Nature Publishing Group (2003)
15. Jaiswal, A., Bach, K.: A Data-Driven Approach for Determining Weights in Global Similarity Functions. In: Case-Based Reasoning Research and Development. pp. 125–139. Springer International Publishing (2019). https://doi.org/10.1007/978-3-030-29249-2_9
16. Jalali, V., Leake, D.: On retention of adaptation rules. In: Case-Based Reasoning Research and Development: 22nd International Conference, ICCBR 2014, Cork, Ireland, September 29, 2014-October 1, 2014. Proceedings 22. pp. 200–214. Springer (2014). https://doi.org/10.1007/978-3-319-11209-1_15
17. Jalali, V., Leake, D., Forouzandehmehr, N.: Ensemble of Adaptations for classification: learning adaptation rules for categorical features. In: Case-Based Reasoning Research and Development: 24th International Conference, ICCBR 2016, Atlanta, GA, USA, October 31-November 2, 2016, Proceedings 24. pp. 186–202. Springer (2016). https://doi.org/10.1007/978-3-319-47096-2_13
18. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: Recommender systems: an introduction. Cambridge University Press (2010)
19. Jeh, G., Widom, J.: Simrank: a measure of structural-context similarity. In: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 538–543 (2002). https://doi.org/10.1145/775047.775126
20. Lenz, M., Burkhard, H.D.: "Case retrieval nets: Basic ideas and extensions". In: Görz, G., Hölldobler, S. (eds.) KI-96: Advances in Artificial Intelligence. pp. 227–239. Springer Berlin Heidelberg, Berlin, Heidelberg (1996)
21. Levary, D., Eckmann, J.P., Moses, E., Tlusty, T.: Loops and Self-Reference in the Construction of Dictionaries. Phys. Rev. X **2**, 031018 (Sep 2012). https://doi.org/10.1103/PhysRevX.2.031018

22. Li, Y., Gao, J., Meng, C., Li, Q., Su, L., Zhao, B., Fan, W., Han, J.: A survey on Truth Discovery. ACM Sigkdd Explorations Newsletter **17**(2), 1–16 (2016)
23. Massie, S., Craw, S., Wiratunga, N.: Complexity Profiling for Informed Case-Base Editing. In: Lecture Notes in Computer Science. pp. 325–339. Springer Berlin Heidelberg (2006). https://doi.org/10.1007/11805816_25
24. Meng, X.L., Van Dyk, D.: The EM algorithm—an old folk-song sung to a fast new tune. Journal of the Royal Statistical Society Series B: Statistical Methodology **59**(3), 511–567 (1997)
25. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781 (2013)
26. Miller, G.A.: WordNet. Communications of the ACM **38**(11), 39–41 (Nov 1995). https://doi.org/10.1145/219717.219748
27. Natarajan, N., Dhillon, I.S., Ravikumar, P.K., Tewari, A.: Learning with noisy labels. Advances in neural information processing systems **26** (2013)
28. P., D., Visweswariah, K., Wiratunga, N., Sani, S.: Two-part segmentation of text documents (Oct 2012). https://doi.org/10.1145/2396761.2396862
29. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web. Tech. rep., Stanford Digital Library Technologies Project (1998)
30. Parsodkar, A.P., P., D., Chakraborti, S.: "Never Judge a Case by Its (Unreliable) Neighbors: Estimating Case Reliability for CBR". In: Keane, M.T., Wiratunga, N. (eds.) Case-Based Reasoning Research and Development. pp. 256–270. Springer International Publishing, Cham (2022)
31. Peat, P..F.D.B..: Glimpsing Reality: Ideas in Physics and the Link to Biology. Routledge (2008)
32. Richter, M.M.: The Knowledge Contained in Similarity Measures. Invited Talk at the First International Conference on Case-Based Reasoning, ICCBR'95, Sesimbra, Portugal (1995)
33. van Rijsbergen, C.J.: The Geometry of Information Retrieval. Cambridge University Press (Aug 2004). https://doi.org/10.1017/cbo9780511543333
34. Schank, R.C.: Dynamic Memory: A Theory of Reminding and Learning in Computers and People. Cambridge University Press, USA (1983)
35. Shekhar, S., Chakraborti, S., Khemani, D.: "Linking Cases Up: An Extension to the Case Retrieval Network", booktitle="Case-Based Reasoning Research and Development". pp. 450–464. Springer International Publishing, Cham (2014)
36. Smyt, B., McKenna, E.: Footprint-Based Retrieval. In: Case-Based Reasoning Research and Development. pp. 343–357. Springer Berlin Heidelberg (1999). https://doi.org/10.1007/3-540-48508-2_25
37. Subramanian, R., Ganesan, D., P, D., Chakraborti, S.: "Towards Richer Realizations of Holographic CBR". In: Sánchez-Ruiz, A.A., Floyd, M.W. (eds.) Case-Based Reasoning Research and Development. pp. 201–215. Springer International Publishing, Cham (2021)
38. Vasudevan, S.R., Chakraborti, S.: Enriching case descriptions using trails in conversational recommenders. In: Case-Based Reasoning Research and Development: 22nd International Conference, ICCBR 2014, Cork, Ireland, September 29, 2014-October 1, 2014. Proceedings 22. pp. 480–494. Springer, Springer International Publishing (2014)
39. Wittgenstein, L., von Wright, G.H., Anscombe, G.E.M.: Notebooks, 1914-1916. Mind **73**(289), 132–141 (1964)