Explainable Case-Based Reasoning: A Survey

Explainable Case-Based Reasoning: A Survey

Jakob M. Schoenborn[1,4,*], **Rosina O. Weber**[2], **David W. Aha**[3],
Jörg Cassens[1], **Klaus-Dieter Althoff**[1,4]
[1]University of Hildesheim, Germany, {schoenb, joerg.cassens}@uni-hildesheim.de
[2]Information Science, College of Computing & Informatics, Drexel University, USA, rw37@drexel.edu
[3]Naval Research Laboratory, Washington, DC USA, david.aha@nrl.navy.mil
[4]German Research Center for Artificial Intelligence (DFKI), Germany, klaus-dieter.althoff@dfki.de

Abstract

Various literature surveys state and confirm a rapid increase in research on explainable artificial intelligence (XAI) in recent years. One possible motivation for this change are legal regulations, including the general data protection regulation (GDPR) but also similar regulations outside of Europe. Another possible reason is the decreasing trust in machine learning systems since both their algorithms and they models they include are often opaque. The desire to retrieve an explanation for a given decision reaches back to the era of expert systems in the 1980s. Decisions made by experts often rely on their stored experiences, yet most XAI approaches cannot provide explanations based on specific experiences because they do not retain them. In contrast, explainable case-based reasoning (XCBR) approaches can provide such explanations, and thus is of interest to XAI researchers. We present a taxonomy of XCBR approaches by categorizing and presenting current methodologies and implementations based on an extensive literature review. This taxonomy can be used by XAI researchers and CBR researchers who are explicitly interested in the generation and use of explanations.

Introduction

Explanations can be viewed from two perspectives. On the one hand, explanations are used to explain the decision making process of a system for a specific decision. On the other hand, the system's decision model itself is the target of the explanation. The need to provide useful and precise explanations on automated decisions is undisputed. A vast number of explainable AI (XAI) approaches and research directions exist, each having their own (dis-)advantages as discussed by Arrieta et al. (2019). Due to the ever rising availability of data, which researchers try to model using data-driven AI, it seems only intuitive to use machine learning - especially deep neural networks - as a mechanism to provide a given decision to the user. However, as these approaches have become more accurate and capable, they have become more complex and opaque. Additionally, these approaches are traditionally designed to solve a specific task, whereas for humans, a decision is much broader (Gunning and Aha 2019; Miller 2017).

We focus on explainable case-based reasoning (XCBR). Case-based reasoning is a methodology that retrieves and adapts experiences to solve a new problem. XCBR aims to support the solution by offering additional explanations to the receiver of the solution (e. g., using visual components such as rainbow boxes or scatter plots).

How do humans solve new problems? Some of us intuitively search in our memories for similar situations, which we might have encountered before - and especially how we acted in those situations, for example, what we used to solve the problem and what outcome has been achieved. By doing so, we think of certain characteristics of the current situation to judge for ourselves whether the previously encountered situation is similar enough to consider whether the way we acted previously is a possible solution to the current problem. Individuals have preferences on which of these characteristics are the most important and weighs these characteristics with respect to their model of the given problem. Using all these elements of a decision making process, we can represent the decision that has been made step by step as one very important factor for the decision making process among many others (Gunning and Aha 2019). We can *explain* the decision. Case-based reasoning (CBR) has been widely noted for its transparency, for example, as the foundation for a Bayesian Case Model as introduced by Kim et al. (2015).

CBR consists of four knowledge containers (Richter 2003): case base, similarity function, vocabulary, and adaptation rules; and four steps (see Fig. 1): retrieve, reuse, revise, and retain (Aamodt and Plaza 1994). The first step, retrieve, takes as input the situation of a user - the current situation. Based on the attributes and their values, which describe the situation, the best n matching cases from the case base will be retrieved, using similarity measurements. During the second step, the retrieved cases will be proposed to be reused (i. e., provided to the user). The proposed cases or the knowledge containers can be adjusted by using adaptation knowledge to provide a tested or repaired case to the user. The user might accept the solution or reject it due to various reasons, for example, the similarity between the input and the retrieved cases is not high enough (Binns et al. 2018). In case of rejection, the next step triggers, revise. The re-

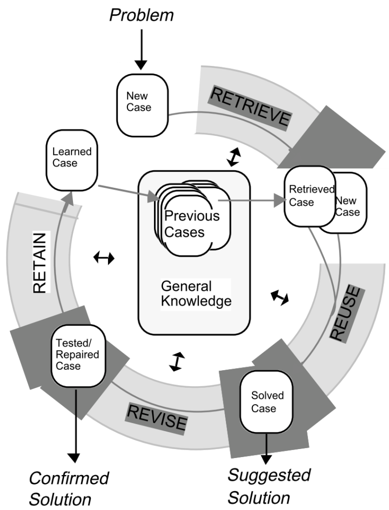

Figure 1: The CBR cycle defined by A. Aamodt and E. Plaza (1994)

vise step typically involves incorporating feedback obtained from testing the proposed solution in a simulation or from a human expert. During the last step, retain, the new case may be retained for further use.

Since the early development of CBR, it has been designed with explainability in mind. Explanation Patterns (XPs) were highly motivated by Schank's Dynamic Memory Organization Packets (Schank 1983) and contained additional elements such as constraints, events, and actors. Specifically, this structure allowed XPs to highlight the reasoning trace as a list of casual forms. Furthermore, this allowed the user to search for conflicting information and in that case to deny an explanation as a direct result (Kass, Leake, and Owens 1986). XPs were used as supplementary textual information in addition to the most similar case and were similar to Schank's canned explanations. For more detailed information, see (Schank 1986; Leake 1992).

The importance of the process of providing an explanation is likely to be an increasing function of decision risk. For example, assuming a person has been diagnosed with breast cancer - is a surgery necessary, or is an endocrine therapy or chemotherapy the better option? To answer this question, and to *explain* the answer, Lamy et al. (2019) presented a visual approach using CBR to support experts by making this important decision, which covers all elements of the decision-making process. Their system then displays the chosen characteristics and corresponding weights to the user, so that an informed decision can be made.

In this survey we present a taxonomy that relates research approaches on explainable CBR (XCBR). In the next section, we will introduce explainable case-based reasoning, and the methodology we used for of our literature review. Interested readers can use our taxonomy to more clearly understand the benefits and limitations of the XCBR approaches we surveyed, as well as to identify opportunities to further the state-of-the-art on this topic.

Explainable Case-Based Reasoning

CBR, and consequently XCBR, can be used in machine learning tasks (for example in supervised learning tasks). However, CBR additionally is a knowledge-based system approach for storing and reusing concrete experiences. One way this has been achieved is by crawling through expert web communities to automatically extract knowledge from textual sources. For example, FEATURE-TAK (Reuss et al. 2016) is an extensible framework initiated in the aircraft domain that has been used to extract information from maintenance engineers to insert relevant information into a CBR system. From a knowledge-based system point of view, this allows many approaches for generating explanations, in contrast to neural networks, which are generally classified as black boxes (Rissland 2006). It may be argued that lazy case-based learning approaches do not derive an abstract model and thus can be treated as a black-box as well. However, similarities and decisions are always based on specific cases (e. g., data points) and how a decision has been retrieved can be reviewed by presenting the impact of each given attribute to the resulting outcome. Additionally, a distinction is made based on the amount of knowledge that is required, typically called *knowledge-light* and *knowledge-intensive* approaches. As an example of the former, the CBR system ProCon (McSherry 2003) highlights both supporting and opposing features in a retrieved case and presents a compiled list of conflicting features to the user, enabling the user to make an informed decision. In contrast, knowledge-intensive approaches include Armengol and Plaza's (2006) use of symbolic explanations of similarities and several efforts on XPs (Kass, Leake, and Owens 1986; Leake 1991; Schank, Kass, and Riesbeck 1989).

CBR is often praised for its rather transparent and realistic model and approach for interpretation and problem solving (Doyle, Tsymbal, and Cunningham 2006; Johs, Lutts, and Weber 2018; Lillehaug 2011). Even without a large training data set, supervised case-based learning approaches can often obtain reasonable results. This is especially helpful since large training data sets might not be accessible due to time, space, or cost constraints. However, it is crucial for a functional CBR system to maintain domain knowledge and to fill its knowledge containers properly (Richter 2003). This maintenance task can be solved by the system developer.

Literature Review Methodology

This section is divided into two parts: We begin by describing the background of our literature review. This includes a definition of explainable artificial intelligence and distinctive characteristics on XCBR. We then describe the methodology we followed to identify the most relevant literature regarding XCBR. We use the results conducted by the literature review to define the taxonomy in the following section.

Background

XAI refers predominantly to explaining the methodologies and results of AI (mostly, machine learning) approaches by increasing the transparency of their processes, and building user trust (Arrieta et al. 2019; Madumal et al. 2018;

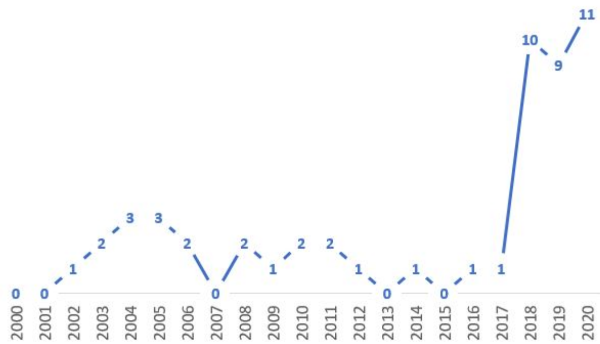

Figure 2: Number of relevant contributions which explicitly discuss XCBR (contents of Fig. 3) by either providing first approaches or theoretical frameworks. The peak starting 2018 is analogous to the rising interest in general XAI as pointed out by Arrieta et al. (2019).

Miller 2017). This desire dates back to 1981: "*Indeed, years ago, explainability was ranked by physicians as the most desirable feature of a clinical decision support system.*" (Teach and Shortliffe 1981). The importance of interaction and consideration of the human as a receiver is pointed out by multiple authors (Arrieta et al. 2019; Miller, Howe, and Sonenberg 2017; Pedreschi et al. 2018). XCBR is a subtopic of XAI. However, there is an important distinction: For the following use of XCBR, we refer to CBR systems that explain their output in contrast to case-based explanations, which describe methods that use CBR to explain other systems. Examples of the latter are the work of Nugent and Cunningham (2005), Weber (2018), and Li et al. (2017). For a general overview on XAI, we refer to the surveys of Adadi and Berrada (2018), Arrieta et al. (2019) and Vilone and Longo (2020). Each provides a comprehensive and structured overview in terms of general concepts, related methods, explanation goals and transparency for multiple approaches to XAI.

We provide a similar overview, but focus on explainable case-based reasoning, which is missing in the current literature (the most recent overview focused on XCBR was written by Sørmo et al. (2005)). Recent publications using XCBR provide examples of its utility, such as data explanation with CBR by Díaz-Agudo et al. (2018), but do not present an overview of current approaches. During the first small hill on further investigating the opportunities of explanations in terms of case-based reasoning (see Fig. 2), Cunningham et al. (2003) discussed the differences among rule-based and case-based explanations for the prediction of blood alcohol level; they reported that users prefer a case-based explanation over rule-based explanation or no explanation at all. The figure indicates that there already has been research on XCBR. With the increased interest on this topic since 2018, we use the opportunity to further investigate and structure the current research on XCBR.

Methodology

To identify relevant literature, we identified commonly used keywords in the XCBR literature, such as "case-based explanation", "explainable case-based reasoning", "XCBR", "explanation patterns" and additionally substituted "explainable" with similar keywords such as "interpretable" and "transparent". Furthermore, we added goals of explanations such as "justification", "transparency", "relevance" and "learning" as introduced by Sørmo et al. (2005) as well as terms of challenges such as "black-box". However, the latter keywords did not lead to additional results compared to the rather direct keywords. We mainly used Google Scholar, IEEExplore, ScienceDirect and arXiv as sources. The query results have been limited to the past 20 years to focus on more recent work.

Additionally, we added an author-based search by visiting the authors' academic profiles and indexing lists on web services such as Google Scholar, Researchgate and DBLP University Trier[1] to identify similar related work by the authors who made contributions to XCBR. This has proven to be helpful to identify a few key researchers who published repeatedly on XCBR during certain time frames.

It is not surprising that most relevant publications on XCBR have been published in the International Conference on Case-Based Reasoning (ICCBR) - including workshops on XCBR and a workshop on XAI held at the International Joint Conference on AI (IJCAI) since 2017.

We identified 52 relevant documents. The decision has been made based on the explanation of their approaches on XCBR by providing a well-structured and well-formulated framework or a concrete implementation, mostly including user studies. These results are shown in Figure 2. In sum, we reviewed 261 documents.

A Taxonomy of Case-Based Explanations

Human-centric explanations

Given that CBR stores concrete events and their corresponding solutions to overcome a problem, it seems intuitive that a most similar retrieved case serves itself as an explanation to the user (Binns et al. 2018). CBR imitates human behavior; humans also think about similar situations whenever we encounter novel situations, which can potentially be solved by adapting a retrieved case's solution. However, solutions vary depending on each person's preferences and thus might not be accepted as an explanation, since the decision has been reached based on data points and feels impersonal (Binns et al. 2018). Indeed, this claim aligns with Adadi et al.'s (2018) observations:

> "*Explanation is, first and foremost, a form of social interaction.* [...] *Based on the conducted analysis, ideas from social science and human behavior are not sufficiently visible in this field*" (Adadi and Berrada 2018, p. 52141f)

Individuals are usually most effective at communicating the motivations and reasons for their decisions by verbal-

[1]https://dblp.uni-trier.de/

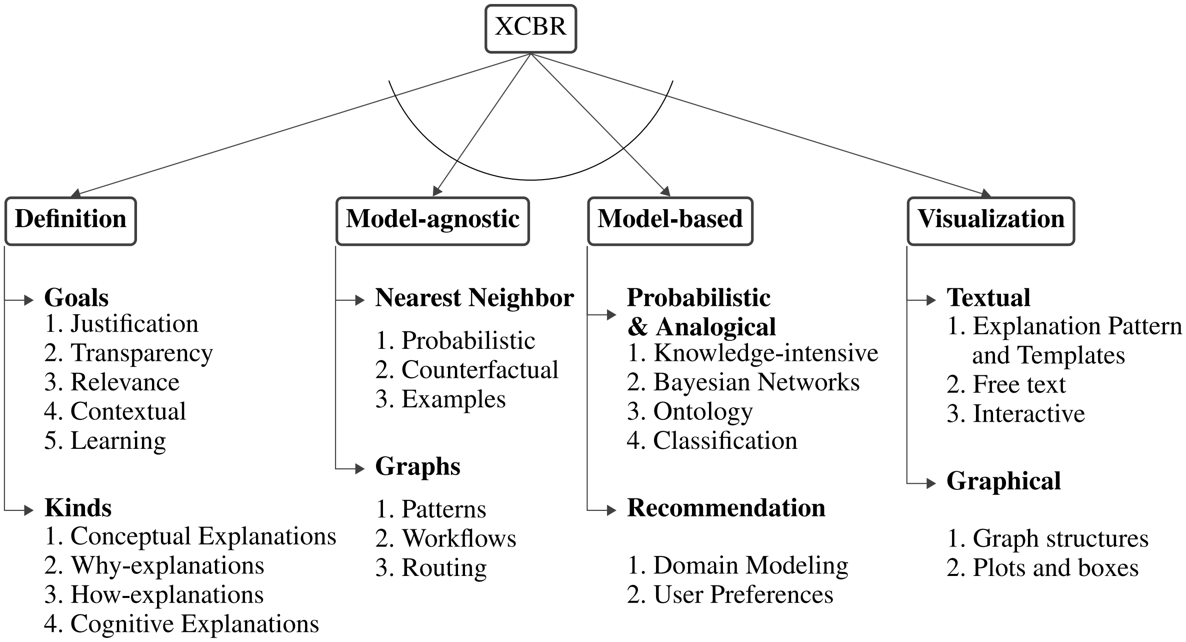

Figure 3: Taxonomy of explanation in XCBR. The categories represent the literature used for Figure 2

izing them. Ehsan and Riedl (2020) introduced "Human-centered Explainable AI" (HCXAI) by focussing on "who" is receiving an explanation. This approach can also be seen in the work of Ribera and Lapedriza (2019), who distinguish explanations between developers, AI researchers, domain experts, and lay users.

In the course of the DARPA Explainable AI (XAI) program, the Oregon State University team is studying models that combine deep adaptive programs, deep reinforcement learning, and techniques for explanation generation (Gunning and Aha 2019). After finding a solution for the mentioned challenges, Holzinger (2018) emphasizes the importance of a simple visualization of an explanation, optimally in at most two dimensions while also focussing on usability, acceptance, and social issues. While these observations are not specific to XCBR, we keep these insights in mind while defining the taxonomy.

Definition of a taxonomy for case-based explanations

We divide our taxonomy into four main areas, inspired by the groupings of Arrieta et al. (2019): Definition, model-agnostic, model-based, and visualization. This decision has been made after investigating the literature on XCBR, which apparently divides up into similar areas as does the XAI literature in general (Adadi and Berrada 2018; Arrieta et al. 2019). Since XCBR is a subset of XAI, decomposing the XCBR literature up into more categories did not seem to be beneficial due to the - at least for now - limited number of publications on XCBR. The existing contributions fan out in multiple directions. Categorizing them into a taxonomy provides a succinct overview on this field of research. Each area covers several sections, such as nearest neighbor, graph structure, and neural network approaches as examples of the model-agnostic area, as these are also often used in general XAI approaches (Arrieta et al. 2019). These categories overlap, especially within *visualization*. For example, the model-agnostic and model-based categories both include approaches with multiple types of visualization techniques. Nevertheless, to increase clarity we include a distinct category for visualization techniques. Figure 3 presents our taxonomy. We next describe the reasoning behind the chosen categories:

Definition. Explanations may differ in their *goal*, such as explaining a XCBR methodology (global explanation) or a specific decision (local explanation). For example, the goals of explanations used in XCBR approaches have included justification (*why the decision has been reached* (Olsson et al. 2014)), Transparency (*how* the decision has been reached (Lillehaug 2011)), Relevance (*which information were relevant for the decision making process* (Roth-Berghofer, Cassens, and Sørmo 2005)), Contextual (*further information on the current situation* (Chari et al. 2020)), and Learning (*teaching the user* (Sørmo, Cassens, and Aamodt 2005)). These goals can be explained by using the questions on the *kinds* of explanations. It is important to consider the goals and kinds of explanations as a foundation for discussion, or otherwise the contribution and increased benefit of the explanation component becomes unclear. In terms of defining goals and kinds of explanations, the XCBR community - and to some extent the XAI community as well - rely on the work and definitions by Roth-Berghofer (2004; 2005), (Sørmo, Cassens, and Aamodt 2005) and (Chari et al. 2020).

Model-agnostic. We distinguish XCBR research as focusing on model-agnostic or model-based approaches, which seems to be the most discriminant factor.

Indeed, providing the most similar case using the nearest neighbor approach aligns with how we intuitively solve problems by retrieving similar situations in our memories (Nosofsky, Sanders, and McDaniel 2018). This does not necessarily require a model to be followed. The model-agnostic nearest neighbor approaches divide into multiple categories: A *probabilistic* approach was presented by (Olsson et al. 2014) among others (e. g., see also (Kern and Virginas 2019) and (Jære, Aamodt, and Skalle 2002)), using machine learning methods such as linear regression as a theoretical foundation to define a similarity measure and the prediction error of retrieving incorrect cases. Keane and Smyth (2020) investigated *counterfactuals*, which convert a nearest neighbor into a nearest unlike neighbor, they used these as contrastive explanations and measured the explanation competence as a function of coverage sets. Counterfactuals or contrastive explanations have also been considered by (Doyle, Tsymbal, and Cunningham 2006), (McSherry 2003), and (Ye et al. 2020). Explanations using *examples* do not need domain knowledge. For example, in the recommendation area, Jorro-Aragoneses et al. (2020) provide textual explanations combined with images of similar movies based on the attributes the user likes.

Graph structures can be used for explanations in different ways. One example is the usage of *patterns* for validation within an application called MetisCBR (Eisenstadt et al. 2018). Their approach has three main steps: pattern recognition, validation, and contextualization, which are processed through an explanation algorithm. The result is a combination of textual explanation combined with a graph structure of the case (based on its attributes), the query, and their similarity. Another example is the use of graphs for explaining *workflows* (Kapetanakis et al. 2010). The workflow event trace is presented and compared to other, similar workflows. Additionally, their system can display which parts of the workflow contribute to the similarity calculation (Kapetanakis et al. 2010). Finally, *routing* is one of the more intuitive applications for using graphs as explanations. Recío-Garcia (2019) and Díaz-Agudo (2019) present alternative traffic navigation routes stored as cases with additional context information (such as weather or traffic conditions).

Model-based. A model can be a helpful tool to provide individual explanations based on the application domain, since information can be presented in a more precise and targeted way. Probabilistic and analogical approaches often require *knowledge-intensive* models. Models allow CBR systems to find similar cases and adapt them to a query, thus reducing the number of cases needed to provide useful explanations. A prominent example is CREEK (Aamodt 2004), which employs substantial general knowledge and whose similarity function is designed to provide an explanation. This has been further developed by (Nikpour, Aamodt, and Bach 2018) using a Bayesian retrieval function in BNCreek.

Recommendation systems tend to focus on the domain model, or the user's behavior. An example for *domain modeling* is described by Ford, Kenny, and Keane (2020); the authors combine case-based explanations within an ANN-CBR twin system and ask the user whether the classification made using the model was correct. Other examples of domain modeling are described by (Sauer, Hundt, and Roth-Berghofer 2012), (Caro-Martínez, Jiménez-Díaz, and Recio-García 2018), and (Paruchuri and Granville 2020). Similar approaches to the beforehand mentioned model-agnostic recommendation approach can be found for *user preferences* by (Martin et al. 2019) and (Recio-García et al. 2019) with a higher focus on the user, using models.

Model-agnostic and model-based approaches are usually (but not exclusively) combined with at least one other methodology to increase the transparency of the decision process, such as representations using graph structures, or visualizations in the form of pictures, diagrams, rainbow-boxes or scatter plots (e. g. Lamy et al. (2019) and Lundberg et al (2018)). For *ontologies*, legal reasoning systems often use them in their decision making process based on previous cases. Ashley et al. (2008) used ontologies in combination with factors, concepts, and principles/policies to assess similar cases. Another typical domain is the medical area. Lamy et al. (2019) presented an approach with a visual component, that depicts case similarity and the modeled weight of the attributes. Lundberg et al.s (2018) approach uses a similar structure and can be easily adjusted to be used as a CBR system.

Visualization. Textual explanations for XCBR are rarely studied. Textual explanations were used by Aamodt (2004) in CREEK to explain the matched cases based on their attributes and their implications on further attributes. A similar, but *interactive* CBR approach was described by McSherry (2003; 2005) to justify recommendations with missing attributes based on a user's preferences. However, recent work seems to use different methods, such as Eisenstadt et al. (2018) methods based on pattern matching, (Machado et al. 2018) reusing past cases and using both, justifications and counterfactuals for providing *free text* explanations as well as Weber et al. (2018; 2019) with a focus on explaining citation recommendations by identifying categories and extending cases., (Gogineni et al. 2018) revisiting *explanation patterns* in the mine clearance domain and (Sizov, Öztürk, and Marsi 2017) uses these as well (which originally have been introduced by (Schank 1983)), report *templates* as explainable cases used by (Díaz-Agudo, Recio-García, and Jiménez-Díaz 2018).

As an example of approaches that rely on visualization for their given explanation, such as using *boxes and plots*, Lamy et al. (2019) presented a therapeutic visual CBR system (see Fig. 4) for breast cancer diagnosis that distinguishes four categories of treatment: surgery, chemotherapy, radiotherapy, and endocrine therapy. Their system combines a quantitative and a qualitative approach. The former is achieved through scatter plots produced by using multidimensional scaling in polar coordinates. While q represents the query, similar cases are scattered around q. The smaller the distance to q, the more similar these cases are. Additionally,

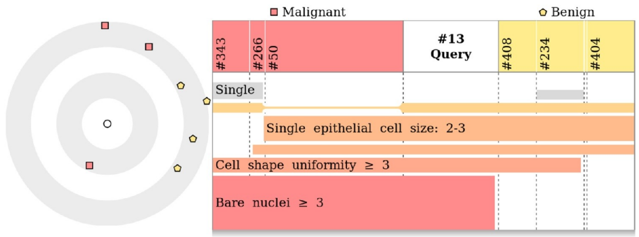

Figure 4: Exemplary visualization of a CBR result, leaving the classification to the user. (left) scatter plots; (right) rainbow boxes (Lamy et al. 2019, p. 51, Figure 6),

shapes are used to assist color-blind users with classifying data points. A qualitative approach has been developed that uses rainbow boxes whose elements represent cases, which are ordered in columns based on their similarity to q. The boxes contain the attributes, which are present in the cases, including their values. The width of the box spans over the cases with this attribute-value pair. The height corresponds to the global similarity of the given attribute. As such, a relatively large box depicts a frequent occurrence of the value in addition to higher emphasis on the given attribute, thus, leading to the possibility for any user to quickly identify the preferences of the CBR system. Furthermore, colors are uniquely associated with each solution class. In case of conflicting attributes (which exist in conflicting treatment categories), the colors are mixed proportionally to their occurrence in a given solution class.

As the authors state, "[...] the interface aims at translating the problem '*which class is* y *and does* q *belong to?*' into a visual problem '*what is the dominant color?*' (Lamy et al. 2019). In Figure 4, the box height of 'bare nuclei' indicates a higher priority of the CBR system on this attribute in comparison, for example, to 'cell shape uniformity'. Since red seems to be the dominant color, the query can be classified as 'malignant'. However, the proposed system does not provide a recommendation regarding which therapy should be chosen, but rather leaves the final choice and thus the final responsibility to the user.

Discussion

XCBR approaches have several advantages and limitations in comparison with other types of XAI approaches. Regarding case-based explanation in general, it can be argued that the user has no insight into the process of how the similarity has been assessed, similar to black-box algorithms. Nevertheless, the case base contains multiple cases, depicting an occurred situation without training data. Generally speaking, XCBR offers a range of opportunities in terms of providing valuable explanations since the approach models the way that individuals recall past, similar situations.

However, XCBR also must confront challenges such as contextual awareness, especially in terms of considering the individuality of each human. A user might refuse to accept an explanation, which has been made based on the experiences of another person - even if that person is very similar regarding demographic data and behavior. This has led researchers to combine CBR with other methodologies to increase explainability (Rissland 2006).

In terms of the taxonomy, this is a first approach to structure the current literature into distinct categories to gain a comprehensive overview on XCBR. We covered the reasoning behind choosing these categories and which kinds of contributions they contain. However, the taxonomy is still broadly defined. This might be revisited in future work by distinguishing between internally and externally used data for the explanation. While we are aware of the technical challenges, we also emphasize the importance of user studies to assess explanation effectiveness, as several researchers have argued (Arrieta et al. 2019; Madumal et al. 2018; Miller 2017; Mueller et al. 2019). Humans are the receivers of an explanation, they should be considered while constructing explanation-aware systems (Miller, Howe, and Sonenberg 2017). The taxonomy helps to solve this task by pointing to different, possible approaches used in the XCBR area.

Conclusions

The challenge remains to provide a satisfactory explanation to the user while considering multiple conditions of the user, such as their current knowledge, emotional commitment, and the context. Machine learning approaches are becoming steadily better in providing appropriate recommendations based on past user behavior and exposure in social networks such as the small world environment (Watts and Strogatz 1998), but providing a proper explanation is still an open challenge due to the black-box nature of most approaches.

We provided an overview on the efforts made in the CBR community to address these challenges by experimenting with different kinds of visualizations, understandings of what a *good* explanation means, models, and audiences. As such, XCBR offers a wide range of methods for generating comprehensible, transparent explanations. However, further investigation is needed on the importance of the receiver of the explanation, especially regarding XCBR. Individuality is an important factor and as such, one might not accept a decision that has been based on the experiences of another, similar person, thus, also not accepting an explanation using a similar approach. To investigate which kinds of explanations are well-accepted among users, large subject studies are still needed and should be a primary focus of future research.

References

Aamodt, A., and Plaza, E. 1994. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Commun.* 7(1):39–59.

Aamodt, A. 2004. Knowledge-intensive case-based reasoning in CREEK. In Funk, P., and Calero, P. A. G., eds., *Advances in Case-Based Reasoning*, 1–15. Springer Berlin Heidelberg.

Adadi, A., and Berrada, M. 2018. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access* 6:pp. 52138–52160.

Armengol, E., and Plaza, E. 2006. Symbolic explanation of similarities in case-based reasoning. *Computers and Artificial Intelligence* 25(2-3):153–171.

Arrieta, A. B.; Díaz-Rodríguez, N.; Ser, J. D.; Bennetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R.; Chatila, R.; and Herrera, F. 2019. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Computing Research Repository (CoRR) arXiv [cs]* abs/1910.10045.

Ashley, K. D. 2008. The case-based reasoning approach: Ontologies for analogical legal argument. In *Approaches to Legal Ontologies: Theories, Domains, Methodologies*. Dordrecht: Springer Netherlands. 99–115.

Binns, R.; Van Kleek, M.; Veale, M.; Lyngs, U.; Zhao, J.; and Shadbolt, N. 2018. “it”s reducing a human being to a percentage’. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI ’18*.

Caro-Martínez, M.; Jiménez-Díaz, G.; and Recio-García, J. A. 2018. A theoretical model of explanations in recommender systems. *XCBR: First Workshop on Case-Based Reasoning for the Explanation of Intelligent Systems — ICCBR 2018: The 26th International Conference on Case-Based Reasoning. At: Stockholm, Sweden* 52–63.

Chari, S.; Gruen, D. M.; Seneviratne, O.; and McGuinness, D. L. 2020. Directions for explainable knowledge-enabled systems. *Computing Research Repository (CoRR) arXiv.* arXiv:2003.07523 [cs.AI].

Cunningham, P.; Doyle, D.; and Loughrey, J. 2003. An evaluation of the usefulness of case-based explanation. In *Proceedings of the 5th International Conference on Case-Based Reasoning: Research and Development*, ICCBR’03, 122–130. Berlin, Heidelberg: Springer-Verlag.

Díaz-Agudo, B.; Caro-Martínez, M.; Recio-García, J. A.; Jorro-Aragoneses, J. L.; and Jiménez-Díaz, G. 2019. Explanation of recommenders using formal concept analysis. In *Case-Based Reasoning Research and Development - 27th International Conference, ICCBR 2019, Otzenhausen, Germany, September 8-12, 2019, Proceedings*, 33–48.

Díaz-Agudo, B.; Recio-García, J. A.; and Jiménez-Díaz, G. 2018. Data explanation with cbr. *XCBR: First Workshop on Case-Based Reasoning for the Explanation of Intelligent Systems — ICCBR 2018: The 26th International Conference on Case-Based Reasoning. At: Stockholm, Sweden* 64–74.

Doyle, D.; Tsymbal, A.; and Cunningham, P. 2006. An evaluation of the usefulness of explanation in a case-based reasoning system for decision support in bronchiolitis treatment. *Comput. Intell.* 22(3-4):269–281.

Ehsan, U., and Riedl, M. O. 2020. Human-centered explainable AI: Towards a reflective sociotechnical approach. *Computing Research Repository (CoRR) arXiv.* arXiv:2002.01092 [cs].

Eisenstadt, V.; Espinoza-Stapelfeld, C.; Mikyas, A.; and Althoff, K.-D. 2018. Explainable distributed case-based support systems: Patterns for enhancement and validation of design recommendations. *In: Case-Based Reasoning Research and Development. International Conference on Case-Based Reasoning (ICCBR-2018) July 9-12 Stockholm Sweden Springer*.

Ford, C.; Kenny, E. M.; and Keane, M. T. 2020. Play MNIST for me! user studies on the effects of post-hoc, example-based explanations & error rates on debugging a deep learning, black-box classifier. *CoRR* abs/2009.06349.

Gogineni, V. R.; Kondrakunta, S.; Molineaux, M.; and Cox, M. T. 2018. Application of case-based explanation to formulate goals in an unpredictable mine clearance domain. *XCBR: First Workshop on Case-Based Reasoning for the Explanation of Intelligent Systems — ICCBR 2018: The 26th International Conference on Case-Based Reasoning. At: Stockholm, Sweden* 42–51.

Gunning, D., and Aha, D. 2019. Darpa’s explainable artificial intelligence (xai) program. *AI Magazine* 40(2):44–58.

Holzinger, A. 2018. From machine learning to explainable AI. In *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)*, 55–66.

Jære, M. D.; Aamodt, A.; and Skalle, P. 2002. Representing temporal knowledge for case-based prediction. In *Advances in Case-Based Reasoning, 6th European Conference, ECCBR 2002 Aberdeen, Scotland, UK, September 4-7, 2002, Proceedings*, 174–188.

Johs, A.; Lutts, M.; and Weber, R. 2018. Measuring explanation quality in XCBR. *XCBR: First Workshop on Case-Based Reasoning for the Explanation of Intelligent Systems — ICCBR 2018: The 26th International Conference on Case-Based Reasoning. At: Stockholm, Sweden*.

Jorro-Aragoneses, J. L.; Caro-Martínez, M.; Díaz-Agudo, B.; and Recio-García, J. A. 2020. A user-centric evaluation to generate case-based explanations using formal concept analysis. In Watson, I., and Weber, R. O., eds., *Case-Based Reasoning Research and Development - 28th International Conference, ICCBR 2020, Salamanca, Spain, June 8-12, 2020, Proceedings*, volume 12311 of *Lecture Notes in Computer Science*, 195–210. Springer.

Kapetanakis, S.; Petridis, M.; Ma, J.; and Bacon, L. 2010. Providing explanations for the intelligent monitoring of business workflows using case-based reasoning. In *Explanation-aware Computing, Papers from the 2010 ECAI Workshop, Lisbon, Portugal, August 16, 2010*, 25–36.

Kass, A.; Leake, D.; and Owens, C. 1986. *SWALE: A Program that Explains*. Lawrence Erlbaum. 232–254.

Keane, M. T., and Smyth, B. 2020. Good counterfactuals and where to find them: A case-based technique for generating counterfactuals for explainable AI (XAI). In Watson, I., and Weber, R. O., eds., *Case-Based Reasoning Research and Development - 28th International Conference, ICCBR 2020, Salamanca, Spain, June 8-12, 2020, Proceedings*, volume 12311 of *Lecture Notes in Computer Science*, 163–178. Springer.

Kern, M., and Virginas, B. 2019. Predicting bid success with a case-based reasoning approach: Explainability and business acceptance. In *Artificial Intelligence XXXVI - 39th SGAI International Conference on Artificial Intelligence, AI 2019, Cambridge, UK, December 17-19, 2019, Proceedings*, 462–467.

Kim, B.; Rudin, C.; and Shah, J. 2015. The Bayesian case model: A generative approach for case-based reasoning and prototype classification.

Lamy, J.; Sekar, B. D.; Guézennec, G.; Bouaud, J.; and Séroussi, B. 2019. Explainable artificial intelligence for breast cancer: A visual case-based reasoning approach. *Artif. Intell. Medicine* 94:42–53.

Leake, D. 1991. An indexing vocabulary for case-based explanation. *Proceedings of the Ninth National Conference on Artificial Intelligence* 10–15.

Leake, D. B. 1992. *Evaluating Explanations: A Content Theory*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Li, O.; Liu, H.; Chen, C.; and Rudin, C. 2017. Deep learning for case-based reasoning through prototypes: A neural network that explains its predictions. *Computing Research Repository (CoRR) arXiv* abs/1710.04806.

Lillehaug, M. B. 2011. Explanation-aware case-based reasoning. Masterthesis at NTNU, Trondheim.

Lundberg, S. M., et al. 2018. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature Biomedical Engineering* 2(10):pp. 749–760.

Machado, N. L. R.; Fontoura, L. M.; Bordini, R. H.; and Silva, L. A. d. L. 2018. Explanation templates for case-based reasoning in collaborative risk management. In *The 30th International Conference on Software Engineering and Knowledge Engineering, Hotel Pullman, Redwood City, California, USA, July 1-3, 2018*, 74–79. KSI Research Inc. and Knowledge Systems Institute Graduate School.

Madumal, P.; Miller, T.; Vetere, F.; and Sonenberg, L. 2018. Towards a grounded dialog model for explainable artificial intelligence. *Computing Research Repository (CoRR) arXiv* abs/1806.08055.

Martin, K.; Liret, A.; Wiratunga, N.; Owusu, G.; and Kern, M. 2019. Developing a catalogue of explainability methods to support expert and non-expert users. In *Artificial Intelligence XXXVI*, 309–324. Cham: Springer International Publishing.

McSherry, D. 2003. Explanation in case-based reasoning: an evidential approach. *Proceedings 8th UK Workshop on Case-Based Reasoning* 47–55.

McSherry, D. 2005. Explanation in recommender systems. *Artif. Intell. Rev.* 24(2):pp. 179–197.

Miller, T.; Howe, P.; and Sonenberg, L. 2017. Explainable AI: beware of inmates running the asylum or: How I learnt to stop worrying and love the social and behavioural sciences. *Computing Research Repository (CoRR) arXiv [cs]* abs/1712.00547.

Miller, T. 2017. Explanation in artificial intelligence: Insights from the social sciences. *Computing Research Repository (CoRR) arXiv [cs]* abs/1706.07269.

Mueller, S. T.; Hoffman, R. R.; Clancey, W. J.; Emrey, A.; and Klein, G. 2019. Explanation in human-ai systems: A literature meta-review, synopsis of key ideas and publications, and bibliography for explainable ai. *Computing Research Repository (CoRR) arXiv [cs]* abs/1902.01876.

Nikpour, H.; Aamodt, A.; and Bach, K. 2018. Bayesian-supported retrieval in BNCreek: A knowledge-intensive case-based reasoning system. In *Case-Based Reasoning Research and Development*, 323–338. Cham: Springer International Publishing.

Nikpour, H.; Aamodt, A.; and Skalle, P. 2016. Diagnosing root causes and generating graphical explanations by integrating temporal causal reasoning and CBR. In *ICCBR Workshops*.

Nosofsky, R. M.; Sanders, C. A.; and McDaniel, M. A. 2018. A formal psychological model of classification applied to natural-science category learning. *Current Directions in Psychological Science* 27(2):129–135.

Nugent, C., and Cunningham, P. 2005. A case-based explanation system for black-box systems. *Artif. Intell. Rev.* 24(2):pp. 163–178.

Olsson, T.; Gillblad, D.; Funk, P.; and Xiong, N. 2014. Case-based reasoning for explaining probabilistic machine learning. In *Case-Based Reasoning Research and Development*, volume 6, 360–374. Cham: Springer International Publishing.

Paruchuri, V., and Granville, B. 2020. A case-based reasoning system for aiding physicians in decision making. *Intelligent Information Management* 12:63–74.

Pedreschi, D.; Giannotti, F.; Guidotti, R.; Monreale, A.; Pappalardo, L.; Ruggieri, S.; and Turini, F. 2018. Open the black box data-driven explanation of black box decision systems. *Computing Research Repository (CoRR) arXiv* abs:1806.09936 [cs]. arXiv: 1806.09936.

Recio-García, J. A.; Díaz-Agudo, B.; Jorro-Aragoneses, J. L.; and Caro-Martínez, M. 2019. SAFEWAY: an explainable context-aware recommender system for safe routes. In *Workshops Proceedings for the Twenty-seventh International Conference on Case-Based Reasoning co-located with the Twenty-seventh International Conference on Case-Based Reasoning (ICCBR 2019), Otzenhausen, Germany, September 8-12, 2019*, 121–132.

Reuss, P.; Stram, R.; Juckenack, C.; Althoff, K.-D.; Henkel, W.; Fischer, D.; and Henning, F. 2016. FEATURE-TAK - framework for extraction, analysis, and transformation of unstructured textual aircraft knowledge. In Goal, A.; Diaz-Agudo, B.; and Roth-Berghofer, T., eds., *Case-based Reasoning in Research and Development. International Conference on Case-Based Reasoning (ICCBR-2016), 24th, October 31-November 2, Atlanta, Georgia, United States*, LNCS. Springer.

Ribera, M., and Lapedriza, À. 2019. Can we do better explanations? A proposal of user-centered explainable AI. In *IUI Workshops*.

Richter, M. M. 2003. Fallbasiertes Schließen (engl. Case-based Reasoning). In *Handbuch der Künstlichen Intelli-*

genz (engl. Handbook of Artificial Intelligence), volume 26. Informatik-Spektrum. 407–430.

Rissland, E. L. 2006. The fun begins with retrieval: Explanation and cbr. In Roth-Berghofer, T.; Göker, M. H.; and Güvenir, H. A., eds., *Advances in Case-Based Reasoning*, 1–8. Berlin, Heidelberg: Springer Berlin Heidelberg.

Roth-Berghofer, T.; Cassens, J.; and Sørmo, F. 2005. Goals and kinds of explanations in case-based reasoning. In *WM 2005: Professional Knowledge Management - Experiences and Visions, Contributions to the 3rd Conference Professional Knowledge Management - Experiences and Visions, April 10-13, 2005, Kaiserslautern, Germany*, 264–268.

Roth-Berghofer, T. 2004. Explanations and case-based reasoning: Foundational issues. In *Advances in Case-Based Reasoning, 7th European Conference, ECCBR 2004, Madrid, Spain, August 30 - September 2, 2004, Proceedings*, 389–403.

Sauer, C. S.; Hundt, A.; and Roth-Berghofer, T. 2012. Explanation-aware design of mobile mycbr-based applications. In *Case-Based Reasoning Research and Development - 20th International Conference, ICCBR 2012, Lyon, France, September 3-6, 2012. Proceedings*, 399–413.

Schank, R. C.; Kass, A.; and Riesbeck, C. K. 1989. *Inside Case-Based Explanation*. USA: L. Erlbaum Associates Inc., 1st edition.

Schank, R. C. 1983. *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*. USA: Cambridge University Press.

Schank, R. C. 1986. *Explanation Patterns: Understanding Mechanical and Creatively*. USA: L. Erlbaum Associates Inc.

Sizov, G.; Öztürk, P.; and Marsi, E. 2017. Let me explain: Adaptation of explanations extracted from incident reports. *AI Commun.* 30:pp. 267–280.

Sørmo, F.; Cassens, J.; and Aamodt, A. 2005. Explanation in case-based reasoning-perspectives and goals. *Artif. Intell. Rev.* 24(2):pp. 109–143.

Teach, R. L., and Shortliffe, E. H. 1981. An analysis of physician attitudes regarding computer-based clinical consultation systems. *Computers and Biomedical Research* 14(6):pp. 542 – 558.

Vilone, G., and Longo, L. 2020. Explainable artificial intelligence: a systematic review.

Watts, D. J., and Strogatz, S. H. 1998. Collective dynamics of ‘small-world’ networks. *Nature* 393(6684):440–442.

Weber, R.; Johs, A.; Li, J.; and Huang, K. 2018. Investigating textual case-based XAI. In *Case-Based Reasoning Research and Development - 26th International Conference, ICCBR 2018, Stockholm, Sweden, July 9-12, 2018, Proceedings*, volume 11156 of *Lecture Notes in Computer Science*, 431–447. Springer.

Weber, R.; Hong, H.; and Goel, P. 2019. Explaining citation recommendations: Abstracts or full texts? *IJCAI 19 XAI Workshop, Macau*.

Ye, X.; Leake, D.; Huibregtse, W.; and Dalkilic, M. 2020. Applying class-to-class Siamese networks to explain classifications with supportive and contrastive cases. In Watson, I., and Weber, R. O., eds., *Case-Based Reasoning Research and Development - 28th International Conference, ICCBR 2020, Salamanca, Spain, June 8-12, 2020, Proceedings*, volume 12311 of *Lecture Notes in Computer Science*, 245–260. Springer.