Utrecht University

Combining A Fortiori Reasoning and a Similarity Measure in Case-Based Reasoning

Author:
Joep Nouwens
(6537618)

Supervised by:
Henry Prakken
Anne-Marie Borg
Floris Bex

A thesis presented for the degree of
MSc Artificial Intelligence

June 26, 2024

Abstract

Over the last years, case-based reasoning has shown to be a promising AI-method in the legal domain, in which a fortiori reasoning is utilized to find similar cases from the past. However, this approach limits the decision making process by not being able to make a prediction for every new case, resulting in a significant number of cases remaining undecided. This thesis discusses the development and evaluation of a newly created case-based reasoning (C-BR) model designed to address this issue, by combining two previously designed models: the aforementioned legal C-BR model and a traditional similarity measure-based C-BR model. Similarity measures are commonly employed in C-BR models to retrieve previous cases in order to solve new problems. The combined approach proposed in this case study includes a fortiori reasoning as well, which involves a formal model of legal reasoning to retrieve similar cases from the past. This combined approach aims to improve the accuracy and reliability of C-BR models. The combined model was tested on a decision making problem of the CBR ('Centraal Bureau Rijvaardigheidsbewijzen': Dutch Central Office of Driving Certification), in which a decision of the fitness to drive of individuals was made based on their health deviations. The performance of the combined C-BR model was compared to models solely utilizing one of the two approaches. While results indicate that the combination of both techniques is a promising approach in C-BR, this model does not outperform the more traditional approach yet, making less human-like decisions than traditional C-BR with a similarity measure. However, this study emphasizes the necessity of further research into the applicability of the integrated model in other domains. Future studies could investigate the potential of the combined C-BR model in datasets that are less complex. This could increase the performance of case-based reasoning models in AI, making these models applicable in even more domains to assist human decision making.

Contents

1 Introduction

1.1 Motivation

In an increasingly complex society, making informed decisions is of great importance for individuals and society as a whole. As humans, we have the capacity to consider various factors, such as personal experiences and external circumstances, when making decisions. However, in numerous situations, it is crucial to make well-informed, objective, and precise decisions that remain unaffected by these experiences and circumstances, as similar situations are otherwise decided differently per individual. This is because individuals often rely on simplifying heuristics when faced with complex information, which can allow personal experiences and external circumstances to influence the decision. Additionally, intuitive or emotional responses also play a significant role in human decision-making [De Martino et al., 2006]. To ensure logical consistency and prevent discrimination or inaccuracy, it is essential to steer clear of human bias when making important choices.

One instance of an area wherein complex decisions have to be made for the safety of our society is road safety. This entire process starts at the institutes for driving, which is the CBR in the Netherlands ('Centraal Bureau Rijvaardigheidsbewijzen': Dutch Central Office of Driving Certification). The CBR determines whether an individual has the ability and the fitness to drive a vehicle on the public roads. Every day, the ability and fitness to drive are determined for a large number of people. Both processes must be done very carefully. Focusing on an individual's fitness to drive, it is important that all possible mental and physical impairments are taken into account to prevent unfit drivers in traffic. Equally important, fit drivers should not be declared unfit incorrectly, which could cause a violation of a person's freedom.

Traditionally, the assessment of driving fitness at the CBR relies on automatic judgments for the most common health deviations, or manual judgments made by medical professionals based on their clinical experience and knowledge. Rules for the most common combinations of health conditions are built into a rule-based system to generate automatic decisions, though more complex combinations of medical deviations have to be considered manually and one at the time. This approach is very time-consuming and could be prone to biases and variations in manual assessments, possibly leading to inefficiency, inconsistency and therefore potential inaccuracy in determining an individual's fitness to drive.

Over the last decades, Artificial Intelligence (AI) models have shown excellent performance in various domains for making predictions and decisions based on patterns and relationships found in data [Morrow and Sormani, 2020], leading to AI-algorithms assisting humans in various domains as in healthcare [Shaheen, 2021]. These AI-algorithms excel at analysing and processing enormous volumes of data, which has led to identifying the potential to develop other models that can assist human decision making [Uddin et al., 2023, Blanco Valencia et al., 2018]. These decisions are often much quicker than the time it would take humans, while the algorithms also ensure that the mental state or personal experiences of the decision-maker do not influence the final decision.

Due to these technological improvements, the next logical progression is that AI-algorithms can support society to maintain safety on the road. Therefore, the CBR aims to develop an AI-algorithm that can assist the medical experts in evaluating an individual's fitness to drive, by advising medical experts by making a decision using an AI-algorithm and showing evidence for this decision. By implementing an AI-algorithm, it becomes possible to develop a system that can analyze an individual's medical conditions and provide an advice of their fitness to drive for the CBR based on previous examples. Implementing such a model could have a significant impact on the decision making process: in this case, the time it normally takes to determine an individual's fitness to drive could be reduced. This is because a driver fitness decision based on previous data can already be available for the medical expert when making the actual final decision. Additionally, the model could reason consistently based on cases from the past. An AI-algorithm could be less prone to human inconsistencies, such as the personal mental states mentioned above.

The application of AI-algorithms comes with some obstacles. Many AI-algorithms, especially machine learning models that train and optimize themselves based on input data and soft coding [El Naqa and Murphy, 2015], lack transparency and explainability because of their complex inner structure [London, 2019]. However, these are crucial elements in Artificial Intelligence [Balasubramaniam et al., 2022] and in the assessment of driver fitness. 'Explainability' is the extent to which the model can produce insights about the reasons for their decisions [Gilpin et al., 2018], and 'transparency' means "the level to which a system provides information about its internal workings or structure" [Tomsett et al., 2018]. The CBR is a semi-governmental organisation, and therefore they must be cautious with the application of AI-models in decision making processes. The adoption of AI could threaten "fundamental rights such as transparency, privacy, autonomy and non-discrimination", as stated in a letter to the government by Van Huffelen, Adriaansens and Dijkgraaf [Van Huffelen et al., 2024]. This means new, unseen problems cannot simply be solved by machine learning models because of their deficit of transparency and explainability. An explainable and transparent model would allow medical professionals of the CBR to understand the reasoning behind the decision-making process [Holzinger et al., 2017], which can lead to experts having more confidence in the decisions made by the model [Samek and Müller, 2019, Gerlings et al., 2020] and generate evidence for the decision at the same time.

One possible solution for this problem is case-based reasoning (C-BR). C-BR is an AI-method that has shown to be a fitting solution for accurate decision making and explainable output when using medical data [Kolodner, 1992, Chattopadhyay et al., 2013, Blanco Valencia et al., 2018, Heindl et al., 1997] and in legal situations [Horty and Bench-Capon, 2012]. C-BR uses a knowledge base (KB) of previously solved problems, which are called 'cases'. These cases are used to determine the solution to a current problem by making a decision based on one or more similar cases in the KB [Kolodner, 1992, Richter and Weber, 2016]. This AI-method is completely transparent, because it only retrieves a similar previous case and does not apply any machine learning [Aamodt and Plaza, 1994]. Furthermore, providing previous cases as an explanation for the current problem makes this method explainable.

Despite being used in legal situations and in the medical field, C-BR has its limitations. C-BR models often use similarity measures to determine the most similar previous problem(s) to a new problem. C-BR models using such a measure base their decision on one or more most similar problems. This sometimes results in not finding the most optimal evidence for a decision from the KB, which will be addressed later. Additionally, the most similar case will in some situations be very close to a case, but for more unique cases much further. Therefore, a similarity measure might provide inconsistent levels of reliability of evidence for its decisions.

For this reason, legal C-BR models sometimes use a fortiori reasoning (AFR) to only return decisions that are derived with certainty based on previous cases in the KB. This legal C-BR model was inspired by the development of CATO by Aleven [Aleven, 1997]. AFR is often used in law and says that if something is true in a certain situation, then it is assumed to be true as well in an even stronger situation [Horty, 2011]. In legal C-BR, all features of the cases are used to determine if one instance is 'stronger' than another. AFR assumes that every feature is for or against a certain decision, and by looking at the values of these features it can be determined if one cases is stronger evidence than a previous case for a decision [Horty and Bench-Capon, 2012]. Applying this reasoning scheme to derive solutions in a C-BR model, might allow for more accurate decision making in a fortiori decisions, as solutions are only inferred if every value is stronger evidence for the decision of a case from the KB. Decisions made with a fortiori reasoning are called 'forced decisions'.

However, C-BR models that use AFR to find a decision a new problem do not always find a forced decision. As the number of features increases, the probability of finding forced decisions decreases, as the possibility exists for more values to be worse evidence for a decision and more combinations of feature values are possible. Consequently, many new cases remain undecided in a fortiori C-BR models.

Thus, the combination of a fortiori C-BR and similarity based C-BR could fill a gap in the scientific field of AI. A possible combination of a fortiori reasoning and similarity measures in C-BR could increase the usefulness of C-BR models in the medical and legal field by possibly returning more accurate decisions than standard C-BR models that solely use similarity measures, by using a fortiori reasoning first. Besides, combining a fortiori reasoning and similarity measures enables to propose more decisions for new problems than previous a fortiori C-BR models, by applying a similarity measure for cases that do not have a forced decision. To obtain such a model, findings of earlier traditional similarity based C-BR approaches are combined with legal a fortiori reasoning C-BR models to create a C-BR model that applies both case retrieval strategies and combines them for optimal decision making.

1.2 Research Question

To summarize, the objective of this research is to investigate whether the integration of a fortiori reasoning with a similarity measure within the context of C-BR will result in more accurate decisions than a C-BR model that employs solely one of these methods. By combining the advantages of both approaches, the accuracy of the decisions of a C-BR model may increase. An increase in accuracy may also enlarge the broader applicability of C-BR methods in AI, and especially in the medical and legal domain, where explainable AI-methods are required because of the complexity of the data and the importance of explainable decisions in these domains.

Furthermore, we aim for an accurate C-BR model that can produce similar decisions to medical experts of the CBR. These C-BR decisions can be used to advise the medical experts of the CBR for new cases, which reduces time in the decision making process. Additionally, a C-BR model utilizes the same domain knowledge of the medical experts, while not being sensitive to human inconsistencies such as being distracted or overlooking similar previous cases in the decision making process. The C-BR model must take into account all health condition categories of individuals in the decision making process. Finally, the output of the C-BR model must be explainable, providing a similar case from the KB as evidence for its decision.

In order to achieve the desired outcomes, it is necessary to address the following questions:

Q1. What are the advantages of combining a fortiori reasoning and a similarity measure in C-BR?
Q2. How can a fortiori reasoning be utilized in the domain of this study?
Q3. How can a similarity measure from the C-BR literature be modified to compare cases in the domain of this study?
Q4. How is the performance of the combined C-BR model measured to compare the model to C-BR models solely utilizing a similarity measure or a fortiori reasoning?

By performing a literature research and employing this material in a case study at the CBR, these questions will be addressed.

1.3 Thesis Structure

The thesis commences with a literature section in Section 2, in which the basics of C-BR is discussed as well as the existing strategies for case retrieval in C-BR models. This section is followed by a detailed explanation of the domain of the case study in Section 3, which is used to provide information about the data that is used for the C-BR model and the process in which the model may be implemented. In Section 4, we describe the weaknesses of current C-BR models and the benefits of combining two case retrieval strategies and the requirements for integrating the two chosen strategies. Section 5 highlights the procedure of comparing the newly developed model to previous C-BR models, and the performance metrics that are used in the comparison. The results, discussed in Section 6 provide the obtained performance scores per model. These are further addressed in Section 7, which is the discussion, in which the limitations of the suggested model are summarized as well. Finally, the conclusion section in Section 8 highlights the findings and discusses further work.

2 Relevant Literature

2.1 AI in the medical domain and its challenges

Artificial Intelligence (AI) has shown to be of great advantage in society, and in medical decision making as well. Already in 1974, computer-based decision making was introduced in the medical domain by the creation of MYCIN [Shortliffe, 1974]. MYCIN was one of the first AI-algorithms used in healthcare and was "designed to assist physicians with the selection of appropriate therapy for patients with bacterial infections" [Shortliffe, 1974]. The introduction of MYCIN led to an increasing amount of research into computer-based decision making. Not long after this, in the 1980s, the first articles were written about general AI-approaches for medical diagnosis [Schwartz et al., 1987]. The first knowledge-based decision-making algorithms were developed [Buchanan, 2005] and the first applications of AI in healthcare emerged: the University of Missouri developed the AI/RHEUM system, which was able to perform rheumatologic diagnosis [Miller, 1986]. In the same decade, SPE was developed, which goal was to interpret results of serum protein electrophoresis data produced by a laboratory instrument. This algorithm contained one of the first forms of AI as well after the development of MYCIN [Kulikowski, 1988].

A few decades later, there has been tremendous progress in the application of AI to the medical field. Recent studies have shown that AI-algorithms are now helping to diagnose patients with possible breast cancer [Uddin et al., 2023], explain medical images using deep learning and Explainable AI [Van der Velden et al., 2022], help with Alzheimer's disease diagnosis and assist in many other aspects of diagnoses in healthcare [Zhang et al., 2022]. In these particular cases, the goal of the models is to provide information to medical experts, so that a well argued decision can be made regarding someone's health condition.

With the rapid growth of AI-algorithms in recent years, many different types of models have been developed to assist in decision making. Each model comes with its own advantages and disadvantages. Machine learning models like neural networks and random forests, which train themselves based on input data, tend to obtain high accuracy scores [Janiesch et al., 2021]. However, decisions by these models cannot be adequately explained because of their complexity. On the other hand, there exist rule-based systems and case-based reasoners. Especially rule-based systems score relatively lower on performance measures like accuracy, though they are really transparent and explainable due to their structure and inner workings. Thus, a trade-off must be made between accuracy and explainability [London, 2019].

As we are dealing with personal health data and the outcome can have an impact on the safety of citizens in the medical domain, it is crucial to make a careful choice of the AI-model. The model must have a high level of confidence in its output, and it is not sufficient to rely solely on the output of an algorithm. Every decision made by an algorithm must be justifiable when decisions are made about the freedom and safety of human lives, which is the case in the medical domain. In other words, Explainable AI becomes crucial to achieving social responsibility. [Gerlings et al., 2020].

Although there have been significant improvements when it comes to explainability of deep learning models in recent studies, decisions by machine learning models still cannot be explained perfectly [Gerlings et al., 2020]. Moreover, some machine learning algorithms have shown discriminatory behavior [Veale and Binns, 2017, Wachter et al., 2020]. For these reasons, standard machine learning models are ruled out for assessing driver fitness, and an alternative solution needs to be considered.

This solution has to be trustworthy, explainable, and transparent [Cutillo et al., 2020]. A model that aligns well with these requirements is Case-Based Reasoning (C-BR) [Watson and Marir, 1994]. C-BR will be discussed further in subsection 2.2.

In addition to the challenge of explainability, AI in the medical field comes with the difficult task of handling the huge quantity and diversity of data that medical systems generate. For AI-systems, the massive volume of data presents significant hurdles, even if it holds the potential for new discoveries and advancements in healthcare. The huge variety in different data types that are stored in medical databases complicates the use of AI-methods such as C-BR. Among other things, types as numbers, truth values, textual values, dates and categorical option values are found in the databases. These data must often be converted into usable data formats so that the data can serve as an input for methods such as C-BR.

So far, the only algorithms that are able to work with a huge variety of different data types are neural networks and other machine learning algorithms, as these models train themselves to work with different types of data [Jiang et al., 2017]. However, as mentioned before, these models cannot be used in certain instances such as (semi-)governmental institutions. Therefore, the basics of standard C-BR models must be adjusted so that more accurate decisions can be made, and C-BR models become more useful.

2.2 Case-Based Reasoning

Case-based reasoning (C-BR) is an AI-approach in which new problems are solved using already solved problems from the past. C-BR uses similar previous problems and uses their solutions to solve new problems [Kolodner, 2014]. These problems are called 'cases', and they are stored in a knowledge base (KB) together with their solution. This solution is also known as a label, and usually refers to a binary output value. C-BR has been a powerful approach in multiple AI-examples, in which decisions were made that had an impact on the freedom of citizens. The method is used in cases which involve the law [Odekerken and Bex, 2020], in healthcare [Kolodner, 1992, Blanco Valencia et al., 2018, Heindl et al., 1997] and in other domains where decisions are made about human lives. C-BR offers several advantages in problem-solving and decision-making processes compared to other AI-methods. This subsection highlights the main components of a C-BR model, discusses key benefits of C-BR and provides practical examples where its application has been successful. Furthermore, it will be discussed how C-BR models can provide explainable evidence for its output.

Unlike many other AI-methods, C-BR does not involve any machine learning [Richter and Aamodt, 2005]. Instead, the model reasons based on its knowledge base (KB). The KB consists of a set of previous cases and other knowledge of the domain, which are both used to find a solution for a new problem. By adapting the current situation to a similar case, the model can return a solution for the current problem. This offers a great advantage, because the decisions made by this AI-model can be explained much better than those made by complex AI-algorithms such as neural networks, support vector machines and others that apply machine learning [Schoenborn et al., 2021]: by providing a solution to the current problem supported by evidence from a similar previous case, the solution to the current case can be clearly justified.

C-BR comes with the following other advantages compared to other AI-methods:

- *Adaptability*: C-BR allows for the adaptation of old solutions to meet new demands. By leveraging past experiences stored as cases, C-BR can effectively address problem instances by modifying existing solutions. For example, in medical diagnosis, C-BR systems can find previously documented similar cases and adapt them to current patients based on their specific requirements [Kolodner, 1992, Heindl et al., 1997].
- *Explanation and interpretation*: C-BR enables the use of old cases to explain new situations. By referring to past cases, C-BR systems can provide explanations for their reasoning and support decision-making processes. For instance, in legal reasoning, C-BR can reason from precedents to interpret new legal situations or create solutions to new problems [Odekerken and Bex, 2020]. This applies for algorithms in medical healthcare as well: in some cases, diagnoses of new patients are based on diagnoses of previous patients with similar conditions [Blanco Valencia et al., 2018].
- *Common-sense reasoning*: C-BR is extensively used in day-to-day common-sense reasoning [El-Sappagh and Elmogy, 2015]. It retrieves past experiences to make informed decisions and solve problems based on similarities with previously encountered situations, much like human reasoning. An example of this is meal planning, where a C-BR system can recommend recipes and meal combinations based on similarities with previously successful meal plans [Kolodner, 1992].

One main formal description of the structure of C-BR models that has been highly accepted, was introduced by Aamodt and Plaza in 1994. They defined the C-BR process as a cycle of four elements, which is a continuous process as long as the C-BR model is used [Aamodt and Plaza, 1994]. The cycle can be observed in Figure 1 and consists of the following elements:

- *Retrieve*: the problem is identified and a similar previous case is retrieved from the knowledge base.
- *Reuse*: the similar case is reused to solve the current problem.
- *Revise*: the proposed solution is revised, often by a human expert.
- *Retain*: finally, the current problem together with the final solution are added to the knowledge base.

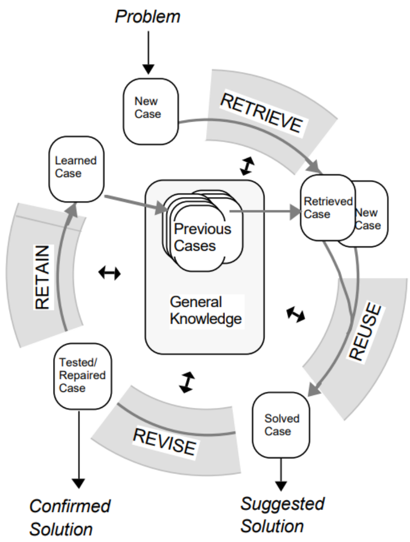

Figure 1: *Cycle of 4Rs in C-BR by Aamodt and Plaza [Aamodt and Plaza, 1994]*

As mentioned above and shown in the center of Figure 1, every C-BR model contains a knowledge base (KB). A C-BR model requires a KB to determine solutions of new cases, which is a collection of past cases together with the corresponding solutions, and further domain knowledge. A 'case' in the KB typically consists of the details of the problem, called 'features', together with its solution [Kolodner, 1992]. The KB has to be managed by human experts who determine which previous cases are sufficient for the KB of the case-based reasoner, and to insert knowledge of the problem domain.

Cases in the KB are stored according to a chosen case representation strategy, which determines how features and other details of cases are stored in the KB. These 'features' are also called 'dimensions' if they have a specific non-binary value. The way a case is stored in the KB of a C-BR model, which is called the 'case representation', contributes greatly to the performance of the model [El-Sappagh and Elmogy, 2015, Finnie and Sun, 2003, Watson and Marir, 1994]. Many different case representation methods exist in C-BR [El-Sappagh and Elmogy, 2015, Bergmann et al., 2005] and the sufficient case representation differs in every problem domain, as every method has its advantages and disadvantages. Existing methods are feature-value pairs, frame-based representations, object oriented representations, textual representations, hierarchical representations and ontologies [Watson and Marir, 1994, El-Sappagh and Elmogy, 2015]. Every method stores different information, as some differ in storage of relevant features and their values, and others contain additional information regarding relations between features. In every problem domain, a specific case representation strategy has to be chosen. The case representation method for this case study is discussed in Section 3.2.1.

The other aspect of the KB is the domain knowledge. The domain knowledge represents all relevant information that is needed for determining the most similar case from the KB. Examples of domain knowledge are information of recognizing exceptions or information for the chosen case retrieval algorithm. Furthermore, the domain knowledge may include an associated value per feature, should a particular feature value of a case be absent. Moreover, a retrieval algorithm of C-BR that utilizes a similarity measure often needs 'weight' values, which represent the importance of each feature of a case to the output. Furthermore, to prevent cases with incorrect input values, a C-BR model sometimes uses boundaries of values to handle incorrect cases immediately before applying all C-BR algorithms.

In the KB of a C-BR model, there should be no constraints between different cases, meaning that cases with equal inputs cannot have different output values. Otherwise, the KB will have two contradicting cases, which both return different output decisions when retrieved from the KB. Moreover, cases containing missing feature values should be removed from the KB or otherwise dealt with appropriately. Additionally, duplicate instances of cases must be removed to prevent the storage of redundant cases. A human expert should be responsible for this task, as this forms the basis for the model and should not contain any errors.

Using the KB, the C-BR model runs through the 'cycle of 4Rs' shown in Figure 1 for each new case. The first element of this cycle is the 'retrieval' process. In this part, the current problem is compared to all other cases stored in the KB to retrieve the solution for the most similar previous case. Different case retrieval methods exist for the comparison to cases in the KB. As mentioned in Section 1.1, the majority of C-BR models employ a 'similarity measure' to determine which cases from the KB are most similar to the current problem. Similarity measures are formulas that calculate the degree of similarity by comparing the features of the current case to the features of a case from the KB. Similarity measures vary based on the type of investigation, since the measure depends on the case representation. In order to find the most similar case in the KB, a similarity measure is applied to every stored case in the KB. One or more cases from the KB that achieve the highest similarity score are then retrieved.

However, some models differ from the standard approach of using a similarity measure to find equivalent cases. For instance, formal models of legal C-BR can be used instead, to identify cases that may restrict the case to one decision. Horty conducted a research in which the basic idea of legal C-BR was introduced. This legal C-BR model used a fortiori reasoning to determine the decision of new cases based on binary feature values of cases in the KB [Horty, 2011]. This approach was later modified by Horty himself to allow for 'dimensions', which are features of cases that account for other data types than just binary values. By applying a fortiori reasoning, the C-BR model will retrieve a similar case when a decision is 'forced' by another case in the KB. A forced decision means that all dimension values of the new case are an even stronger evidence for a decision than a case that is stored in the KB, from which only one decision follows. This type of reasoning could increase the accuracy and reliability of the model, as the model will make decisions that are consistent with decisions of previous cases. Moreover, it will only retrieve a similar case if the model is sure of a decision. The downside of a fortiori reasoning as retrieval method is that many new cases remain unanswered, especially when the number of features per case is high. It is therefore necessary to consider carefully which case retrieval method is most appropriate for a problem situation.

The second process of C-BR in Figure 1 is the 'reuse' process. In many situations, a new case has to be adapted to meet the requirements of the previously stored cases. This step is especially important if similarity measures have been used to find equivalent cases. Most of the retrieved cases from the KB are not entirely equal to the current problem, coming with different feature values. These differences have to be made clear to create a correct suggested solution, so that the C-BR user can observe the degree of similarity between the cases.

In the 'revise' procedure, the suggested solution is checked for correctness. These decisions of the C-BR model on new cases have to be manually checked. Human decision support is necessary in this process, as each decision has to be revised by people who have enough understanding of the decision making process. Even though the algorithm can be trusted when its accuracy is high enough, it remains important to check every decision carefully when the retrieved similar case is not entirely similar to the new problem. Otherwise, the experts may suffer from the 'control problem'. The control problem states that a human expert may overestimate the accuracy of the case-based reasoner, leading to a failure to detect errors made by the model. In order to prevent this, human experts should be kept in a constant loop of evaluating new cases [Odekerken and Bex, 2020]. If the suggested solution is correct, the C-BR model will use the proposed solution in the next step. If this is not the case, the human expert corrects the decision. Additionally, if the retrieved case from the KB is incorrectly decided, this case must be reconsidered or removed. This prevents incorrect decisions due to inconsistencies in the KB and increases the accuracy of the C-BR model.

The last step in the cycle is to 'retain' the current case. After obtaining the correct solution of the human expert, the algorithm performs the final step. First, the C-BR algorithm determines if the current case already exists in the KB. If not, the current case is added to the KB in the correct case representation. Thereby, C-BR model has acquired knowledge by incorporating new information into the knowledge base in the form of a solution to a new problem.

To compare the new case to other cases, a similarity measure or another case-comparing strategy must be programmed which is suitable for the data that is used. This strategy is referred to as the 'case retrieval strategy'. This case retrieval strategy must be applied to compare all the cases in the KB to a new problem. Finally, the solution to the current case, based on the most similar case from the past, must be adapted based on the differences between the cases. Once the solution to the current problem is verified by the human expert, the algorithm stores the new case with its solution in the KB.

As mentioned in Section 1.1, most models use a similarity measure to compare new cases to previous cases in the KB in the 'retrieval' phase. Even though this is the most simple solution for this process, it remains the question if this method is the most accurate strategy. However, the advantage of applying a similarity measure is that it always allows for a most similar case in the KB. Before being able to discuss the existing retrieval strategies and developing a possibly better strategy, it is of importance to further discuss the basic foundations of the two existing case retrieval methods. These methods can then be combined to possibly increase the accuracy of C-BR models.

2.3 A Fortiori Reasoning

In the case retrieval stage of the C-BR cycle, one possible solution to retrieve a similar case from the KB is using a fortiori reasoning (AFR). AFR is a formal model of legal reasoning. The reasoning model refers to a deduction that says that if a claim is true in one instance, it must be even more true in another where the evidence is stronger. Refraining from this reasoning model would result in inconsistent reasoning. AFR is based on a familiar human practice [Canavotto and Horty, 2022]. This argumentation strategy has been adopted from legal reasoning procedures in court [Horty, 2011], and has been applied in previous case-based reasoners [Odekerken and Bex, 2020]. The main advantages of this type of reasoning is that it systematically prevents wrong decisions in a KB by only allowing decisions for cases that contain at least as strong evidence as a previous case with this decision. Assuming that the model reasons on correct situations from the past, applying AFR will only generate decisions if every aspect of a new situation is at least as strong evidence for that decision. This type of reasoning is applied in the legal domain [Horty, 2011].

In order to get a better understanding of AFR, an example will be provided in Equation 1, based on the current case study for the CBR.

1. **The new individual** is more **healthy** than **an individual in the knowledge base**
2. **and the individual from the knowledge base** is **healthy** enough to be **fit to drive** (1)
3. therefore, all the more, **the new individual** is **healthy** enough to be **fit to drive**

As follows from this example, this type of reasoning ensures that a decision only follows if it is logically consistent with all previous decisions. Any individual that complies with the first inference cannot be declared otherwise than stated in the third inference. To apply these inference rules in a C-BR model, we briefly summarize a formalized theory of AFR that was created by Horty [Horty, 2011]. Horty described his formalized theory of AFR as the 'Result Model', which described AFR for cases that only consisted of binary features. This formal theory was further elaborated later into the 'Dimensional Result Model', allowing for more values than only binary values. A formal summary of this model was described by Van Woerkom et al. [Van Woerkom et al., 2023], which is described below.

2.3.1 Dimensional Result Model

The Dimensional Result Model states that the decision of every case is determined based on dimensions, which are features of a case that have a numeric or categorical value. Every dimension has its set of possible values, which depend on the type of the dimension. Dimensions are denoted with lowercase letters, and the set of possible values of dimension $d \in D$ is denoted by V_d.

A *fact situation* X contains all dimensions and the corresponding values of a case. It is a combination of possible valuations of the set of dimensions D, defined by the function $X : d \to V_d$, such that $X(d) \in V_d$ for every dimension $d \in D$. The definition of a case is given in Definition 2.1 below.

Definition 2.1 (Case)**.** A case c consists of a fact situation X and a solution s, denoted $c = \langle X, s\rangle$, where X contains all features and their values, and s equals one of the two possible solutions.

Cases are decided for one of two possible solutions of the binary classification problem, denoted by π and δ. It is assumed that every dimension $d \in D$ has its own *preference relation*, which describes the preference of every value towards one of the solutions. The formal definition of a preference relation is denoted in Definition 2.2.

Definition 2.2 (Preference Relation)**.** A preference relation $\preceq$ of dimension $d \in D$ is denoted by $d : v \preceq_\pi w$ for v, $w \in V_d$, where we say w prefers outcome π relative to v. Equivalently, v prefers outcome δ relative to w, denoted by $w \preceq_\delta v$, where $\preceq_\delta$ is the opposite of $\preceq_\pi$. A preference relation is required to be a partial order.

Cases are decided based on possible constraints in the KB. A new fact situation X_{new} which values all prefer outcome s relative to the values of the fact situation of another case C_{old} in the KB with solution $s \in \{\pi, \delta\}$, cannot receive another solution than the solution of C_{old}. This is called a *forced decision*. A formal decision of a 'forced decision' is given in Definition 2.3.

Definition 2.3 (Forced Decision)**.** A decision of a fact situation X for side s is forced by a case base KB, denoted $KB, X \models s$, if there is a case $(Y, s) \in KB$ such that $Y(d) \preceq_s X(d)$ for all $d \in D$.

To clarify this theory of the Dimensional Result Model, an example will be given. Consider a binary classification scenario in the medical field, where the goal is to predict whether or not a patient has a bacterial infection. The two solutions are $\{0, 1\}$, where 0 equals *no infection* and 1 stands for *infection*. Assume the classification is based on the following dimensions $fe, rn, st, bt, ot \in D$ and their ranges:

– fe: presence of a fever, with values {*false*, *true*}
– rn: presence of a runny nose, with values {*false*, *true*}
– st: presence of a sore throat, with values {*false*, *true*}
– bt: body temperature, with values between 36 and 42°C
– ot: outside temperature with values between -20 and 40°C

The preference relations of the dimensions $d \in D$ are defined as follows:

– fe: $false \preceq_1 true$
– rn: $false \preceq_1 true$
– st: $false \preceq_1 true$
– bt: $36 \preceq_1 37 \preceq_1 ... \preceq_1 41 \preceq_1 42$
– ot: $40 \preceq_1 39 \preceq_1 ... \preceq_1 -19 \preceq_1 -20$

Suppose there exists a KB that consists of the set $\{C_{old_1}, C_{old_2}, C_{old_3}\}$, and the solution of fact situation of C_{new} has to be decided. Table 1 contains all relevant information of these four cases.

	C_{old_1}	C_{old_2}	C_{old_3}	C_{new}
fe	*true*	$false$	*true*	*true*
st	*true*	$false$	$false$	$false$
rn	*true*	*true*	*true*	*true*
bt	39	37	38	38
ot	5	15	10	2
s	1	0	1	?

Table 1: Values of cases $C_{old_1}, C_{old_2}, C_{old_3}$ and C_{new}

Now we want to determine whether a new case, denoted C_{new}, is predicted to have a bacterial infection. For every case in the KB, it will be verified if a decision is forced for C_{new}. To ascertain whether C_{old_1} forces a decision for C_{new}, this is accomplished by inserting the values of C_{old_1} and C_{new} from Table 1 in Definition 2.3. Suppose X is represents the fact situation of a case. Filling in all variables in the definition of a forced decision, it is given that case C_{old_1} forces decision 1 if $X_{C_{old_1}}(d) \preceq_1 X_{C_{new}}(d)$ for all $d \in D$. For every dimension, this statement is checked below.

- $C_{old_1}(fe) \preceq_1 C_{new}(fe):$ $true \preceq_1 true$ is correct,
- $C_{old_1}(st) \preceq_1 C_{new}(st):$ $true \preceq_1 false$ is incorrect: $false$ prefers 0 relative to $true$,
- $C_{old_1}(rn) \preceq_1 C_{new}(rn):$ $true \preceq_1 true$ is correct,
- $C_{old_1}(bt) \preceq_1 C_{new}(bt):$ $39 \preceq_1 38$ is incorrect: 38 prefers 0 relative to 39,
- $C_{old_1}(ot) \preceq_1 C_{new}(ot):$ $5 \preceq_1 2$ is correct.

Since the value of C_{old_1} for the dimensions st and bt does not force decision 1 for C_{new}, it can be concluded that C_{new} is not forced by C_{old_1}. Similar validations are done for C_{old_2} below.

- $C_{old_2}(fe) \preceq_0 C_{new}(fe):$ $false \preceq_0 true$ is incorrect: $true$ prefers 1 relative to $false$,
- $C_{old_2}(st) \preceq_0 C_{new}(st):$ $false \preceq_0 false$ is correct,
- $C_{old_2}(rn) \preceq_0 C_{new}(rn):$ $true \preceq_0 true$ is correct,
- $C_{old_2}(bt) \preceq_0 C_{new}(bt):$ $37 \preceq_0 38$ is incorrect: 38 prefers 1 relative to 37,
- $C_{old_2}(ot) \preceq_0 C_{new}(ot):$ $15 \preceq_0 2$ is incorrect: 2 prefers 0 relative to 15.

Case C_{old_2} does not force decision 0 for case C_{new} because of the values of dimensions fe, bt and ot. Finally, we investigate C_{old_3}:

- $C_{old_3}(fe) \preceq_1 C_{new}(fe):$ $true \preceq_1 true$ is correct,
- $C_{old_3}(st) \preceq_1 C_{new}(st):$ $false \preceq_1 false$ is correct,
- $C_{old_3}(rn) \preceq_1 C_{new}(rn):$ $true \preceq_1 true$ is correct,
- $C_{old_3}(bt) \preceq_1 C_{new}(bt):$ $38 \preceq_1 38$ is correct,
- $C_{old_3}(ot) \preceq_1 C_{new}(ot):$ $10 \preceq_1 2$ is correct.

As follows from case C_{old_3}, a decision of 1 is forced by the KB because it holds that 'for all $d \in D$, $X_{C_{old_3}}(d) \preceq_1 X_{C_{new}}(d)$'. A fortiori, the model's decision should be 1.

Further requirements for the implementation of AFR will be discussed in Section 4.2.

2.4 Similarity Measures

Another approach is a more common method of finding similar cases in C-BR, which is through the use of a similarity measure. C-BR models that apply a similarity measure in the case retrieval phase compare a new case to all cases in the KB using a mathematical formula. This formula calculates the degree of similarity between two cases by comparing all present (values of) features of the new case and a case from the KB.

In order to find the most similar case, the case from the KB with the highest value of similarity is returned. The application of a similarity measure to find evidence is called 'nearest neighbors' [Liao et al., 1998, Cunningham, 2008]. This approach was described in detail by Finnie and Sun [Finnie and Sun, 2002]. In a paper about similarity in C-BR, Finnie and Sun denoted the relation $R(x, y, u, v)$, which represented that 'x and y are at least as similar as u and v. Using this relation, they conducted the formal definition of 'nearest neighbors', which is denoted in Formula 2.

$$NN(x, z) \Leftrightarrow \forall y R(x, z, x, y) \tag{2}$$

The formula states that z is the nearest neighbor if and only if for every y, z is at least as similar to x as y to x. Applying this formula to cases in a C-BR model, we obtain "case z is the nearest neighbor of new case x if and only if z is at least as similar to x as any other case y is to case x" [Finnie and Sun, 2002]. Even though this perfectly captures the idea of a nearest neighbor, Finnie and Sun do not describe a general similarity measure based on the principle of nearest neighbors. Translating the definition of a nearest neighbor to a mathematical formula, we obtain Formula 3. This equation states that the most similar case to a new case is the case in the KB that returns the highest similarity value, which is equivalent to the lowest dissimilarity value.

$$sim_n \;=\; \underset{c_i \in C}{\operatorname{argmax}}\; SIM(c_n, c_i) \;=\; \underset{c_i \in C}{\operatorname{argmin}}\; DISS(c_n, c_i) \tag{3}$$

where

$$\begin{aligned} sim_p &= \text{the most similar case in the KB,} \\ SIM &= \text{the chosen similarity measure,} \\ DISS &= \text{the chosen dissimilarity measure, which is the inverse of } SIM, \\ c_i &= \text{case } i \text{ in the KB, with i ranging from the first to the last case in the KB,} \\ c_n &= \text{the new case.} \end{aligned}$$

A wide variety of formulas exist that represent a similarity measure. Measures can differ in the level of detail, domain knowledge and type of data used in a C-BR model. The most simple similarity measures are based on binary features. An example is given in Formula 4.

$$\textit{Jaccard Similarity:}\;\; SIM(C_i, C_j) = \frac{|C_i \cap C_j|}{|C_i \cup C_j|} \tag{4}$$

In this formula, $|C_i \cap C_j|$ represents the number of matching features of cases C_i and C_j, and $|C_i \cup C_j|$ stands for the total number of features of the cases. The measure simply returns the fraction of common features between two cases. However, obviously, these similarity measures are often inadequate for complex datasets such as medical data, for the exact same reason as the extension and modification of AFR Result Model that only allowed for binary features: as many problems contain dimensional features as well, similarity measures were generated that took into account other types of feature values. This caused the development of more advanced similarity measures that accounted for dimensions. Two commonly adapted similarity measures in C-BR that base their calculation on dimensions were described by Lance and Williams [Lance and Williams, 1966], which are shown in Equation 5 and 6. Note that these formulas calculate the degree of dissimilarity instead of similarity. Thus, the lower the dissimilarity, the higher the similarity between two cases.

$$Clark: \quad DISS(C_i, C_j) = \sum \frac{|C_{ik} - C_{jk}|^2}{|C_{ik} + C_{jk}|^2} \tag{5}$$

$$Canberra: \quad DISS(C_i, C_j) = \sum \frac{|C_{ik} - C_{jk}|}{|C_{ik} + C_{jk}|} \tag{6}$$

In Formulas 5 and 6, C_i and C_j represent the selected cases, and k denotes the k^{th} feature of the case. These measures were developed to calculate a dissimilarity value between two cases that consist of numerical features, where differences between feature values are penalized more in the measure of Clark. For every value, the difference between the two numbers is divided by the addition of the numbers. Having similar values, this measure returns 0, and values far away will return values closer to 1.

Although these measures provide a more detailed analysis than the simple similarity measure in Formula 4, the measures in Formulas 5 and 6 do not suffice in most C-BR models. In many C-BR applications, certain features or dimensions have a greater influence on the final decision than others. In situations where domain knowledge is unavailable the measures described above can be a great solution, though knowledge of the data is often available before creating a C-BR model. As a consequence, many similarity measures accommodate for feature weights in their calculation.

Feature weights are values that determine the importance of a dimension compared to other dimensions. These weights are stored as extra domain knowledge in the KB, with one value being stored for each dimension. Feature weights can be computed manually, though these are often determined through the use of machine learning [Park et al., 2004, Yeow et al., 2014]. By the addition of weights per dimension to the similarity formula, differences between values of important dimensions contribute more to the output compared to less important dimensions. While this addition already complicates the similarity measure compared to the first measure, not every domain only contains numerical or binary features. Consequently, Castro et al. developed a similarity measure so that it accommodated for feature values, and the existence of categorical dimensions as well [Castro et al., 2009]. This measure was simplified by Kang et al. [Kang et al., 2013]. The dissimilarity measure based on this similarity measure can be found in Formula 7, which is a 'generalized weighted dissimilarity measure', according to Núñez et al. [Núnez et al., 2004]. This measure is denoted, as this is the simplest version of the measure created by Castro et al.

$$DISS(C_i, C_j) = \frac{\sum_{k=1}^{n} w_k \cdot d(C_{ik}, C_{jk})}{\sum_{k=1}^{n} w_k} \tag{7}$$

In this formula, C_i and C_j represent the two cases, w_k is the weight assigned to feature k and $d(C_{ik}, C_{jk})$ is the dissimilarity degree between the value of attribute k in the two cases C_i and C_j. The value of the dissimilarity $d(C_{ik}, C_{jk})$ can be computed as follows according to a modified version of Castro et al. [Castro et al., 2009]:

$$d(C_{ik}, C_{jk}) = \begin{cases} \frac{|C_{ik}-C_{jk}|}{|F_k^{max}-F_k^{min}|}, & \text{if } F_k \text{ (the } k^{th} \text{ feature) is numeric.} \\ 0, & \text{if } C_{ik} = C_{jk}\text{, and 1, otherwise (if } F_k \text{ is nominal).} \end{cases}$$

For numerical features, the dissimilarity is computed by calculating the difference between the values of C_{ik} and C_{jk}, divided by the minimum and maximum value of the dimension. This scales the distance of this dimension to an output value between 0 and 1, leading to an equal contribution to the output for every feature (without feature weights). This formula differs the most from unweighted measures because of the variable w_k, which implies weight specification for each feature in the case representation. Although determining weights for each specific health category is very complex and time consuming, weights can be important if one feature weight is definitely more important than others.

However, in complex domains with many features, it might be difficult to create a sufficient weight assignment that perfectly represents the contribution of every dimension to the output without using any machine learning. As the CBR does not allow for machine learning in a C-BR model, utilizing this similarity measure may not be sufficient. The definitive choice of a suitable similarity measure for this case study will be further discussed in Section 4. Furthermore, it will be discussed in Section 4 what alternative approach of case retrieval will be applied in this case study to possibly increase the reliability and accuracy of C-BR models.

3 Centraal Bureau Rijvaardigheidsbewijzen (CBR)

This section highlights the details of the current driver fitness assignment of the CBR, for which the C-BR model will be developed. In addition, the data used in this driver fitness assignment of the CBR will be discussed. This will give an overview of the data that that will be employed for the case base and the domain knowledge of the C-BR model. Finally, this section will discuss potential barriers and difficulties with the CBR data which need to be addressed in order to use this data in the model.

3.1 Driver Fitness and the Existing Assessment Process of the CBR

Before diving into decisions regarding the implementation of a C-BR model, it is important to clarify and understand the case study in which a C-BR model will be implemented. Therefore, the current process of driver fitness determination of the CBR will be discussed, as well as some recurrent terms at the driving institute. The CBR is a perfect example for situations in which large medical data collections are stored, and where previous cases could serve as a solution for new cases. Therefore, the driving fitness process of the CBR will serve as a practical example of the implementation of a new retrieval strategy in C-BR.

The main goal of implementing a case-based reasoning model at the CBR is to advise the medical experts of the CBR regarding the 'driver fitness' or 'fitness to drive' of an individual. An individual's 'fitness to drive' means the mental and physical capability of participating in motorized traffic in the interest of road safety. Thus, being 'fit to drive' means that a person is healthy enough to participate in traffic safely based on their mental and physical health. Driver fitness encompasses medical fitness, practical driving fitness in the case of a disability or impairment, and driving proficiency in terms of behavior. Examples of people who are determined not fit to drive are people with certain (combinations of) mental or physical disabilities, people who violated the law, or people with specific diseases that might hinder the ability to drive a vehicle.

In order to have a clearer understanding of the driver fitness assessment, a description of the current medical process at CBR will be presented. An individual has several reasons to undergo a medical assessment when it comes to their fitness to drive. The most common scenario involves someone applying for a driver's licence for the first time. Alternatively, it may be that someone is renewing their existing licence, undergoing a re-evaluation of their licence as a result of a traffic violation or a re-evaluation as part of the objections procedure. An objections procedure is followed if an individual believes that they have been unfairly rejected, which causes a reconsideration of the fitness to drive decision.

In every previously mentioned situation, the individual has to fill in a health declaration. The CBR has composed specific questions for this health declaration to determine the mental and physical condition of an individual (see Appendix 9.5). The health declaration consists of 19 closed questions, subdivided into the following 15 categories:

- Arms and legs;
- Vision and eyes;
- Diabetes Mellitus;
- Impaired kidney function;
- Respiratory or blood disorder;
- Transplantation;
- Heart and blood vessels;
- Neurological disorder;
- Epilepsy;
- Drowsiness and unconsciousness;
- Meniere's disease;
- Behavioral disorder and psychiatric diagnosis;
- Addictive substance abuse;
- Medications;
- Additional complaints or conditions.

Based on the 19 questions on the health declaration, the CBR can already partly determine whether an individual is entirely physically and mentally healthy or not. Answering all 19 questions with 'No' automatically means the individual is declared fit to drive.

Alternatively, when one or more questions are answered with 'Yes', it means that the individual is not completely 'healthy'. However, they can still be considered fit to drive if their medical conditions do not significantly impact driving safety to a point where it becomes too dangerous. To determine the impact of the medical deviations of an individual, the CBR makes further specifications of every deviation. The CBR uses manually created rules to determine an individual's 'natures' ('aard-en') in case of one or more 'Yes' answered questions. These natures represent specific categories for mental or physical deviations.

Based on the answers on the health declaration and the corresponding natures, an individual may be referred to a medical examiner or specialist. The situation is discussed with this specialist, and an additional questionnaire with relevant questions is filled out. These answers provide a clearer understanding of the individual's medical condition, allowing for the addition of 'severities' to the natures. These severities stand for a specification of the mental or physical deviation. Each nature and severity comes with a specific value as well, which complements the nature and severity with additional information about the degree of the deviation.

An example of one existing nature (category) in the CBR database and the possible severities (sub-categories) can be found in Table 2, in which the possible severities and values are shown for the nature 'VISUS'. VISUS is the main category for eye defects. This main category has several subcategories that specify whether the right (_VOD) or left (_VOS) eye was measured, or both (_VODS). Additionally, different categories are created to specify whether the value of the eye measurement was with (_MC) or without (_ZC) a correction. The 'Value' column indicates the value of the eye measurement in this combination of natures and severities, which can range from 0.0 (representing complete blindness) to 3.0 (representing exceptional eyesight).

Nature	Severity	Value (min - max)
VISUS	VISUS_VOS_MC	0.0 - 3.0
VISUS	VISUS_VOD_MC	0.0 - 3.0
VISUS	VISUS_VODS_MC	0.0 - 3.0
VISUS	VISUS_VOS_ZC	0.0 - 3.0
VISUS	VISUS_VOD_ZC	0.0 - 3.0
VISUS	VISUS_VODS_ZC	0.0 - 3.0

Table 2: Example of the categorization of Natures and Severities, with a certain value

In the example provided above, all nature and severity combinations come with a numerical value. However, values can be of other types as well, such as boolean values, text values, date values or categorical values. Boolean values often represent the presence of a certain deviation, while numerical values provide information about the seriousness of the deviation. Date values stand for the beginning of a certain disease, such as diabetes, or the latest occurrence of a certain medical deviation, such as a hypo in case of having diabetes. The exact meaning of a value is specified in the nature and severity names. An example of a categorical value can be found in the nature 'FUNCTIEBEP_BOVENSTE_LEDEMA', which represents the category for a defect of the upper body. The severities determine in which part of the upper body the deviation is located. Possible severities in the nature FUNCTIEBEP_BOVENSTE_LEDEMA are severities for deviations of the arm(s), shoulder(s), finger(s), elbow(s) and others. The value per severity determines the side of the deviation, which could be 'left', 'right' or 'left/right' in this category. However, the range of the possible values differs for every combination of nature and severity.

Once the individual has completed the process of examination by medical professionals, their total list of natures and severities is complete. Further details of this data are discussed in Section 3.2.1. The combinations of natures and severities are stored in the data system of the CBR, and the only process left is determining the final decision on the individual's fitness to drive.

The driving license category is an additional factor that influences the decision on driver fitness. Since there exist different driver's licenses for different types of vehicles and the CBR assesses driver fitness for each category separately, specific procedures must be followed for each license category. It is possible for an individual to be deemed fit to drive a standard car, even though this individual is decided not to be fit enough to drive a truck. Per driver category a different final decision can be made. All various categories are shown in Table 3. However, due to limited time and resources, this case study will focus on the driving category B, since most individuals are evaluated for the driver's license of a car.

Category	Vehicle license
A	moped or scooter
B	car
BE	car with trailer
C	truck
CE	truck with trailer
D	bus
DE	bus with trailer
T	tractor

Table 3: Driving categories

Depending on the driving category and list of natures and severities, different procedures are followed. The CBR already has specific rules for certain combinations of natures and severities. For these cases, an automatic decision is made by the implemented rules in their system. For other combinations, a medical expert has to conduct a decision manually. The process is time-consuming due to the necessity of a medical expert to review all nature-severity combinations and their values. Therefore, the C-BR model could be employed to reduce the time required for this process. The manual process of driver fitness evaluation could be sped up by the implementation of the C-BR model if the C-BR model is able to reproduce decisions of medical experts well enough. The C-BR model could already give an advice for the decision of a new case based on the present natures and severities and their corresponding values, by returning a similar case and its decision from the KB. In case of an implementation of the model, the medical expert has yet to make the final decision, but an advice will already have been generated by the model. In addition, an explanation will be given by returning the case(s) from the KB that led to the decision.

3.2 CBR Data

3.2.1 Data Storage

The required data from the CBR has many similarities with other medical data collections. The main aspect of the required data for the driving fitness evaluation consists of a table with all medical health deviations per individual. The total dataset of the CBR can be divided into three different parts. The first table contains all cases together with the corresponding decision date, which represents the day on which a driver fitness decision was made by a medical expert of the CBR. Furthermore, the driving category of the case is specified in this table. Table 4 shows an example of this data.

ID	Decision Date	Category
x	1-1-2024	B

Table 4: Case information data for case x

All final decisions per individual and per driving category are stored in a table containing the case process ID, the driving category and the final decision. An example can be observed below. This table shows that case ID x was decided to be 'Geschikt' ('fit to drive') for the driving categories B, BE and T. Alternatively, an individual can be decided 'Ongeschikt' (unfit) or 'Afzien' (refrain from deciding 'fit'). In addition, two additional properties are given to the decision, namely 'shorter validity' and 'EU codes'. The first denotes whether the 'fit to drive' decision should have a shorter validity period than the standard of 10 years. The second, 'EU codes', contains a value if the driver is fit to drive only under certain conditions, such as a modification of the vehicle.

ID	Category	Decision	Shorter_Validity	EU_code
x	B	Geschikt	-	-
x	BE	Geschikt	-	-
x	T	Geschikt	-	-

Table 5: Decision data for case x

The most important aspect of the dataset for the C-BR model is the table that describes an individual's medical deviations. Every row of this table contains a present health deviation of an individual. This table can be seen as the relevant features of a case that lead to the final decision in the decision making process by the CBR, as the medical experts of the CBR base their final decision on the present health deviations per individual. Consequently, these health deviation categories and their values can be used in the C-BR model as dimensions. For each individual, there may be several lines in this table. Table 6 shows the contents of the table for an example case x of the dataset. Every row in the table contains a main category of a deviation ('nature'), the subcategory of the deviation ('severity') and the associated 'value' per nature-severity combination.

ID	Nature	Severity	Value
x	DM	DM_BEGINDATUM_GETAL	2018
x	DM	DM_NIET_INSULINE_HYPO_-	TRUE
x	VISUS	VISUS_VOD_ZC	0.8
x	VISUS	VISUS_VOS_ZC	0.8

Table 6: Medical features for an example case

The data of example case x is described in the table above, which contains four combinations of natures and severities. The first row shows the first year of having diabetes for individual of case x, which was in 2018. Additionally, it shows extra information about the disease in the second row. The third and fourth row represent eye measures for the right and left eye. Based on this information, the fitness to drive of individual x is evaluated. In total, the dataset of the CBR contains 399 valid combinations of natures and severities. Based on the present natures and severities and their value, medical experts determine if an individual is fit to drive.

Since the combinations of the natures and severities are seen as the features of a case, and every feature comes with its own value, the data is already structured as feature-value pair case representations. This makes the data easy to implement in C-BR models. Moreover, similarity measures can easily apply calculations on feature-value pairs and AFR needs this type of case representation as well, which can be concluded from Definition 2.1. Finally, the feature-value approach provides an intuitive and understandable case representation. This methodology simply represents cases as a collection of features, each of which is associated with a corresponding value. Such a representation promotes ease of interpretation and manipulation by users, regardless of their technical knowledge. By adopting the feature-value approach, the C-BR model becomes more accessible to a wider range of users, especially for medical experts that most likely do not have programming backgrounds. These medical experts can easily understand cases by analyzing the feature-value pairs, thereby enabling streamlined decision making and problem solving. Thus, feature-value pairs will suffice as the case representation strategy in the CBR domain.

Per case in the KB of the CBR, every combination of a specific nature and severity can only occur once per case ID. Most case IDs are associated with only a few rows, since the majority of the individuals only have a couple of deviations or less. The database stores only the relevant deviations, so the absence of a nature-severity combination indicates the absence of the corresponding deviation for the individual.

Unfortunately, data entry errors are inevitable in systems that rely on manual user input for data population. This is especially true when data fields are governed by predefined rules dictating acceptable values, such as specific ranges for the 'value' field or requirements for further specification of subcategories like 'severities' in the CBR data. Deviations from these guidelines are common. Instances may occur where values fall outside certain ranges or users neglect to complete required specifications, represented by the 'severity' and the 'value' column. These inconsistencies lead to undesirable consequences: as a result, these errors compromise the reliability of the dataset for the usage in C-BR models. Implementing strong validation mechanisms is crucial to identify such errors, ensuring the integrity and usefulness of the dataset.

3.2.2 Data Limitations

As denoted above, manual data entry leads to mistakes in the dataset. Moreover, data in the medical field already often contain missing values or other errors [Goldberg et al., 2008]. These errors can hinder the AI-methods that are applied to the data by forcing them to make incorrect decisions [Tschandl, 2021]. In AFR, incorrect values can result in incorrect forced decision making, which decreases the accuracy of the reasoning approach. In a C-BR model that uses a similarity measure, errors could cause inaccurate decisions as well.

The CBR data contains four main types of errors that can pose challenges for C-BR models. The most common issue is the presence of empty cells in critical columns such as 'severity' and 'value'. Medical experts simply forget to enter the corresponding data value in the 'severity' or the 'value' cells. Furthermore, medical experts sometimes fail to fill in the correct data type in the 'value' field for a nature and severity combination. The final, but even more problematic error, is the presence of two identical cases with different output values. C-BR models prefer one output value per input combination, because multiple outputs would make it impossible to determine the correct decision for a new, similar case. By eliminating cases with errors, a C-BR model will be able to find similar cases correctly in most situations.

One final limitation of the complex data of the CBR is the presence of textual fields. Medical experts may write out specifications of an individual's deviation, which cannot be described using only numerical and categorical data values. These specifications are stored in specific nature-severity combinations for textual fields. However, since a fortiori reasoning requires numerical or categorical values, these textual fields cannot be handled in C-BR models that use a fortiori reasoning. Therefore, cases that contain these textual fields must be removed from the KB in the a fortiori reasoning process. The exact way incorrect dimensions and values are handled in this case study is specified in Section 5.2.

4 Combining A Fortiori Reasoning and Similarity Measures in C-BR

In this section, the relevant literature will be applied to the research domain of the CBR, which shows the possible advantages of combining AFR and a similarity measure in C-BR. The chapter starts with a substantiation of the proposed combination, after which the application of the two methods is further discussed.

4.1 Advantages of combining AFR and a similarity measure

While C-BR applications traditionally use similarity measures to find a similar case in the KB [Kolodner, 1992], a fortiori reasoning has shown to be a successful case retrieval method as well [Horty, 2011, Odekerken and Bex, 2020]. However, there have been no studies investigating the performance of C-BR models that combine the methods of a fortiori reasoning and similarity measures in the retrieval process. Therefore, this thesis proposes a new insight in the field of AI, especially for C-BR models, by combining the two retrieval strategies. The following subsections will highlight the potential advantages of combining the two possible retrieval methods.

4.1.1 C-BR without a Similarity Measure

First, we discuss C-BR that solely applies a similarity measure in the case retrieval phase. AFR in C-BR is used to find forced decisions for new cases, as discussed in Section 2.3. When a C-BR model only employs AFR to find similar cases, there will be a vast number of cases without any final decision. This is because for many cases no forced decisions are found, which happens in situations where for every case, the valuation one or more dimensions of a case ensure that a case is not better evidence for a previous decision. The probability of this increases with the number of dimensions that a C-BR model includes in its decision making process. Alternatively, AFR could find forced decisions for both sides, which can occur as well in AFR C-BR models. This happens when decisions for cases in the KB contradict each other because of their dimension values. An evidence will be found for both sides, which causes a contradiction. Details of two-sided forced decisions will be addressed in Section 4.2.1 as well.

If many cases remain undecided due to no forced decisions or double-sided forced decisions, this renders the C-BR model useless since a fortiori reasoning only decides cases that are forced for one decision. When a C-BR model considers more dimensions with a wide range of values, relying solely on a fortiori reasoning becomes less effective because of lower chances of one-sided forced decisions by other cases. The C-BR model will only decide for a small number of cases, and the other cases will all have to be decided manually (without any advice of the C-BR model).

Besides, many dimensions of the CBR domain contain non-categorical textual values. Dimensions that are not categorical or numerical cannot be handled by AFR algorithms. Therefore, cases containing such dimensions remain unanswered as well by AFR models.

By implementing a similarity measure for cases that did not receive a one-sided forced decision, or for cases that contained dimensions insufficient for the use of AFR, the C-BR model can still make a decision for these cases. This could potentially make the model more effective than a C-BR model that only uses a fortiori reasoning. The new cases without forced decisions will then have a decision by the C-BR model as well, not leaving them undecided.

4.1.2 C-BR without A Fortiori Reasoning

On the other hand, C-BR that does not use a fortiori reasoning to find similar cases might be a sub-optimal use of C-BR as well, although the explanation for this side is more complex. The main argument for this side, is that decisions by C-BR models with a fortiori reasoning could be more accurate when using AFR, because similarity measures might find the incorrect decision in certain situations. In contrast to C-BR applying a similarity measure, a fortiori C-BR only decides for cases that have better evidence for a decision than a previous case in the KB with that decision. Therefore, every decision by a C-BR model that applies AFR is consistent with a previous case from the KB. These statements will be explained using the example given in Table 7.

Feature	Old Case 1	Old Case 2	New Case
$VISUS\_VOS\_ZC$	1.0	0.8	0.8
$VISUS\_VOD\_ZC$	1.0	0.8	0.8
$VISUS\_VODS\_ZC$	1.0	0.8	0.8
$VISUS\_VOS\_MC$	1.0	0.8	0.8
$VISUS\_VOD\_MC$	1.0	0.8	0.8
$VISUS\_VODS\_MC$	0.3	0.5	0.2
Decision	Unfit	Fit	?

Table 7: C-BR example

Suppose these features are used to determine whether someone is fit to drive. Here, every feature starting with 'VISUS' refers to a specific eye measure, and the value of these features can range from 0 to 3. 0 is the worst eye measure possible, and 3 means perfect eyesight. The final decision in this C-BR model can either be 'Fit' (to drive) or 'Unfit'.

Using a similarity measure that takes into account the numerical values of the features, in which all dimensions contribute equally to the final decision, the C-BR model will find that Old Case 2 is the most similar case to the New Case in the KB, because all values of the New Case are closer to Old Case 2 than to Old Case 1.

Alternatively, suppose that AFR was applied for making the driver fitness decision. Suppose we have the following preference relation: a high value of every 'VISUS' feature has a preference for the decision 'Fit', and a low value of 'VISUS' feature has a preference for the decision 'Unfit'. With this preference relation, we can use a fortiori reasoning. Since every value of the New Case is equally good or even worse than the values of Old Case 1, it follows a fortiori that the New Case must be 'unfit to drive' as well. Here, we see that both approaches base their answers on two different cases from the KB. In this example, the decision by AFR is more accurate, because Old Case 1 in the KB forces decision 'Unfit'.

The example above describes the main disadvantage of only using similarity measures in this C-BR model. In C-BR models that only use a similarity measure, we are not entirely sure of the decision making of the New Case, since the New Case often completely matches one of the cases in the KB. Thus, the addition of a fortiori reasoning to C-BR models that use similarity measures could be of great added value to possibly obtain more accurate decisions in C-BR models.

4.1.3 Advantages of Combining the Retrieval Strategies

To sum up, combining a fortiori reasoning and similarity measures in the retrieval process of C-BR could improve the accuracy of the model. AFR can be used in the initial case retrieval procedure to decide for a new case. In the event that the case contains dimensions that are unsuitable for AFR, when no forced decisions are found, or when forced decisions for both sides are found, the C-BR model will apply a similarity measure to make a decision for a new case. The combination of these two C-BR approaches could possibly improve the performance of C-BR models.

4.2 A Fortiori Application

The following subsection highlights the requirements for AFR in the domain of the CBR and describes the practical implementation for AFR.

4.2.1 Requirements for AFR

Integrating the Dimensional Model of AFR from Section 2.3.1 within a C-BR model to compare situations requires several critical components. The following aspects must be considered:

- Binary solution framework;
- Preference relations for every dimension $d \in D$;
- Missing value treatment;
- Removal of cases containing unmanageable dimensions;
- Transforming non-linear numerical dimensions;
- Preferable: 'consistent' case base.

First of all, the Dimensional Model of AFR requires a binary solution framework, which allows each situation to be classified into one of two possible distinct outcomes. Thus, the CBR data must be transformed into a binary classification problem to develop a C-BR for driver fitness evaluation. In Section 3.2.1, it was shown that in our case study each individual case was labeled with one of the three decisions 'Geschikt' (fit), 'Ongeschikt' (unfit) or 'Afzien' (refrain from deciding 'fit'). In order to be able to reason with a fortiori reasoning, two of these classifications must be added together to obtain a binary classification algorithm. This will be discussed further in Section 5.2.

Furthermore, a preference relation must be created for every dimension of the cases, to guarantee that the model reasons correctly based on the previous cases, its dimensions and the dimension values. These preference relations determine the positive or negative contribution of every dimension to the final decision of the model. For every data type, the preference relation is constructed differently.

The creation of preference relations is fairly straightforward for binary features, as one value always prefers one decision and the other value prefers the other decision. Using domain knowledge, preference relations for these binary features can be easily generated. However, for other data types, the preference relation is a little more complicated. These types will be discussed below.

Features with numerical values often come with a linear relation. For example, features like *eyesight percentage* or *measured alcohol percentage* have some linear relation with the outcome variable determining someone's fitness to drive. For such features with a linear relation, higher values have a preference for one decision, while lower values prefer the other decision.

While most numerical features represent a linear relationship that is easy to capture in a preference relation, not all features have this linearity: for instance, consider the features with nature 'CORRECTIESTERKTE' in the CBR case study, which represents the *strength of the correction* of an individual's eye lenses and ranges from -20 to 20. This dimension fails to satisfy a linear relationship, as both excessively low and high values are undesirable. Values around 0 are considered 'good' values, as they positively impact the 'fitness to drive' decision. Conversely, values above and below this range have a negative influence on the final decision compared to 0. While certain applications of AFR allow for nonlinear dimensions, a vast number of AFR algorithms do not have the ability to develop preference relations for such dimensions. To establish preference relations for such dimensions, values of these features must be scaled or transformed to ensure a clear linear relationship with the outcome. One solution would be to convert this numerical feature to a binary feature by using rules with boundaries. For this specific feature, values around 0 would be decided with 'low correction', while other values above a certain limit (-10 and +10, for example) would obtain 'high correction'. Using such rules and boundaries, preference relations can be created for nonlinear numerical features by converting the numerical feature a binary feature. Another solution would be to apply a mathematical formula to the original value, which will then capture linearity. For this feature, using the absolute value of the original value already creates a linear relationship, as high values would be worse than lower values.

For categorical features, a similar conversion method can be applied. Many categorical variables are ordinal. Ordinal categorical features, such as the severity of a certain deviation (*light, middle, heavy*), act like numerical features in the way that they capture a clear ordering of the feature values. By substituting a categorical value with a numerical value ($light = 0$, $middle = 0.5$ and $heavy = 1$), the ordinal relation can be captured in a numerical feature. This ordinal relation is similar to a linear relation, which creates the opportunity of developing preference relations for such features. Therefore, ordinal categorical features can be converted to numerical features using numerical values, to be able to create preference relations.

However, categorical features can be nominal as well, which are categorical features that lack a linear relation. For these features, it is not possible to construct a preference relation. AFR can only be applied using cases that have equal values for these features. Unfortunately, a fortiori reasoning cannot be applied to cases that have different values in the present nominal features. One example of a nominal feature in the CBR domain is the feature 'HERKEURING__HERKEU_BESL', with its possible values 'CODES', 'KG', 'KG_EN_CODES' and 'ONGESCHIKT'. This feature denotes that the current case is created because of a reconsideration of a previous decision, and the value represents the reason of this reconsideration. For these values, a preference relation cannot be generated, as they do not have a preference for a decision. From this follows that this feature has to be removed from the dataset, and cases containing this feature have to be removed as well.

Another requirement for using AFR is that empty values should be handled correctly. As discussed in Section 3.2.2, incorrect data entry often occurs in (medical) datasets, including the appearance of empty values. AFR does not allow missing values in the comparison of dimensions. Therefore, cases containing missing values must be removed, or an alternative approach to dealing with these missing values must be devised. For a fortiori reasoning, the most logical solution would be to delete cases that contain an empty feature value. However, the disadvantage of this solution is that the KB of the CBR contains empty values in every row as only a few nature-severity combinations are present per case. Consequently, removing cases with empty column values would make the application of a fortiori reasoning in this case study impossible, as every case contains empty dimensions. While there has been a development of missing value treatment techniques by adding average feature values or the most common feature values [Batista and Monard, 2003], applying such techniques could lead to inaccurate AFR forced decisions. Therefore, using the domain knowledge of the medical experts of the CBR, the empty values can be filled with the most favorable value of every dimension for a positive fitness to drive decision, based on its preference relation. This is because the presence of a nature-severity combination indicates a present health deviation of the individual, which means that the absence of the dimension would be equivalent to having the best possible value for this dimension. An example in the CBR domain would be the feature COGNITIE__COGNITIE_MMSE, which represents results of a cognitive test. The absence of this nature-severity combination indicates that this cognitive test was not executed and a cognitive deviation for this individual is not present. The best possible value for this test is 30, indicating high cognitive ability, and the lowest value is 0, indicating no cognitive abilities at all. Consequently, filling empty values with the best possible value will have a similar influence on the fitness to drive as an empty value, as value of 30 prefers a positive fitness to drive decision compared to lower values. This means inserting the value with highest preference for a positive decision is seen as a sufficient empty value treatment method in this domain to be able to use AFR.

The fourth aspect of AFR requirements is the removal of unmanageable dimensions. AFR allows for numerical and categorical features, but features without specified ranges or categories cannot be handled by the basics of AFR. Even if AFR allowed for textual values, these textual values are often misinterpreted by AI-algorithms because of ambiguity [Poesio, 1995], and because extracting useful information from text and comparing this information is very time-expensive and difficult for an AI-model. For that reason, dimensions that simply contain written text as values cannot be handled by AFR. Cases containing such dimensions must be removed from C-BR, as they are required to be solved differently.

Finally, having a 'consistent' KB is preferred to ensure the accuracy of the decisions of the model. A KB is *consistent* if there are no initial constraints in the dataset. Definition 4.1 formally explains a consistent KB.

Definition 4.1 (Consistent KB)**.** A KB is consistent iff the KB does not contain two cases $c = \langle X, s \rangle$ and $c' = \langle X', \bar{s} \rangle$ such that $X \leq_s X'$ [Peters et al., 2023]

In other words, there cannot exist a case c with outcome s which dimension values all prefer the other outcome $\bar{s}$ more than the dimensions of case c' that has outcome $\bar{s}$. Without a logically consistent KB, the model will find forced decisions for both sides for certain cases, which makes the model less able to make decisions *a fortiori* correctly.

Unfortunately, it is almost impossible to generate a KB that is completely consistent. This is because human decisions are often sensitive to personal influences or their mental states, which sometimes causes a different decision per individual. Therefore, it is important that medical experts carefully investigate every generated decision by the C-BR model and to remove cases from the KB that cause incorrect decisions. These incorrect cases can be found when the C-BR model returns such cases as evidence for a new decision. Deleting these cases will improve future decisions by the C-BR model.

4.2.2 Application of AFR

AFR can be applied in C-BR by implementing the programmed model of Van Woerkom [Van Woerkom, 2023]. Van Woerkom developed this version of a fortiori C-BR using the Z3 Solver of Microsoft [Bjørner et al., 2019].

The Satisfiability Modulo Theories (SMT) Solver Z3 from Microsoft, abbreviated to Z3, is used to efficiently solve SMT problems. SMT Solvers are problem solvers for logical formulas [Bjørner et al., 2019]. Z3 takes propositional formulas as input and is able to validate whether a formula is true, and under which variable assignments a formula could be true. Van Woerkom implemented this Z3 Solver for the development of an 'AFCBR' (a fortiori case-based reasoning) model [Van Woerkom, 2023], a generalized C-BR model that applies a fortiori reasoning in Python.

The program of Van Woerkom first initializes the KB of the C-BR model. Based on this KB, the preference relations are automatically generated, which can be manually altered in the event of inappropriate preference settings. Following the generation of the propositional formulas for cases in the KB, the AFCBR model combines all formulas per decision to obtain one long propositional formula per decision. These formulas can be employed to check for forced decisions.

To obtain usable preference relations for the Z3 model, the program converts every dimension value to an integer value. Numerical dimensions are simply stored using 'ascending' or 'descending' to denote the preference of a dimension. In terms of Definition 2.2, ascending means: for $v, w \in V_d$, if $w > v$, then w prefers 1 relative to v. Consequently, descending stands for the opposite, substituting '1' by '0'. Dimensions that represent a binary variable or categorical values are given an integer value as well, so that every value in the domain of a dimension $d \in D$ gets its own value. To clarify: for values $x, y, z \in V_d$, if x prefers 1 relative to y and z, and y prefers 1 relative to z, then $z = 0$, $y = 1$ and $x = 2$. The preference relation of this dimension will be: $(0)z \leq (1)y \leq (2)x$, with the numbers in brackets indicating the value used by the model for the categorical features.

The employment of AFR in the newly developed C-BR model will be further specified in Section 5. The generated relations and other data preparations will be discussed in that section.

4.3 Suitable Similarity Measure

If no forced decision is found in the 'a fortiori' process of the model, one or more most similar cases are selected from the KB using a similarity measure. In contrast to AFR, the application of a similarity measure does not require a KB in which every dimension has a non-empty value, as empty dimensions can be dealt with in the measure itself based on the developed similarity measure. Thus, empty values will be dealt with in the similarity measure instead of filling the KB, as this allows for less computation time because empty values can be compared more easily. Additionally, while AFR does not allow for textual fields in the reasoning model, a similarity measure can account for non-categorical textual dimensions. The presence of a filled textual value has to be taken into account even when textual fields do not match entirely. For example, suppose two cases have a present CHRONISCH_HARTFALEN__CHRONISCH_HART_OVERIG dimension, denoting a textual explanation of the presence of chronic heart failures. Even though the values of these cases differ, this still makes the cases more similar than two cases where only one has the chronic heart failure dimension. Therefore, textual fields are taken into account in the similarity measure, and cases containing such textual fields are not removed from the KB for the similarity measure.

Moreover, as the CBR data consists of many different types of dimensions and existing methods do not account for empty values, date values and other textual values than categorical dimension values, an existing similarity measure will be modified to allow for these extra types and the empty values.

A traditional similarity measure that already takes into account two different dimension types is given in the leftmost mathematical formula of Equation 8. This measure is based on the similarity measure of Castro et al. in Equation 7, which takes into account every dimension for numerical and categorical values [Castro et al., 2009]. However, since the contribution of every dimension towards the output is unknown, the weights of the dimensions are all set to 1. Replacing the weight per dimension w_k with 1 for every dimension k, we obtain the final formula in Equation 8. In this measure, C_i and C_j are the cases that are compared for dissimilarity and n is the number of dimensions.

$$DISSIM(C_i, C_j) = \frac{\sum_{k=0}^{n} w_k \cdot d(C_{ik}, C_{jk})}{\sum_{k=1}^{n} w_k} = \frac{\sum_{k=1}^{n} 1 \cdot d(C_{ik}, C_{jk})}{\sum_{k=1}^{n} 1} = \frac{\sum_{k=1}^{n} d(C_{ik}, C_{jk})}{n} \quad (8)$$

In the article by Castro et al. [Núnez et al., 2004], $d(C_{ik}, C_{jk})$ calculated the similarity between two values. As denoted in Section 2.4, the dissimilarity between two values is equal to the following formula:

$$d(C_{ik}, C_{jk}) = \begin{cases} \frac{|C_{ik} - C_{jk}|}{|F_k^{max} - F_k^{min}|}, & \text{if } F_k \text{ (the } k^{th} \text{ feature) is numeric.} \\ 0, & \text{if } C_{ik} = C_{jk}, \text{ and 1 otherwise (if } F_k \text{ is nominal).} \end{cases}$$

In this case study, however, this definition is not sufficient. Firstly, the columns of the KB contain many empty values. An empty value represents the absence of a health deviation, which is important information for a case. This makes it important for the formula to take into account. An advantage of this formula is that the output of $d(C_{ik}, C_{jk})$ is constrained to a range between 0 and 1, which means every dimension's dissimilarity is a value between 0 and 1. Therefore, the aim is to develop a measure that maintains this range of values, as it is essential that the contribution of every feature must be equal to the final similarity value. Formula 9 denotes the final dissimilarity value of values C_{ik} and C_{jk}, which accounts for different types of values and empty cells.

$$d(C_{ik}, C_{jk}) = \begin{cases} 0, & \text{if } C_{ik} = C_{jk} \text{ (or both are empty);} \\ 1, & \text{if } C_{ik} \text{ is empty or } C_{jk} \text{ is empty and not both;} \\ 0.75, & \text{if } C_{ik} \text{ and } C_{jk} \text{ are text values (and not equal);} \\ \frac{|C_{ik} - C_{jk}|}{|C_{ik} + C_{jk}|}, & \text{if } C_{ik} \text{ and } C_{jk} \text{ are numbers and } |C_{ik} - C_{jk}| \leq |C_{ik} + C_{jk}|; \\ date(C_{ik}, C_{jk}), & \text{if } C_{ik} \text{ and } C_{jk} \text{ are date values and } date(C_{ik}, C_{jk}) \leq 1; \\ 1, & \text{otherwise.} \end{cases} \quad (9)$$

where

$$date(C_{ik}, C_{jk}) = datedifference\_in\_days(C_{ik}, C_{jk})/10000 \quad (10)$$

Formula 9 calculates the dissimilarity between every dimension for cases C_i and C_j and always returns a dissimilarity value between 0 and 1. First, if two values are equal or if they are both empty, the dissimilarity will obviously be equal to 0, as both dimensions are equal.

If only one of the dimensions is empty, the dissimilarity will be equal to 1, since the nature-severity combination is only present in one of the two cases. If both values are text values (and not equal), the dissimilarity value will be 0.75, as the presence of dimension values shows a more similarity between the two cases than the absence of one of the two values.

If C_{ik} and C_{jk} are numerical values, the dissimilarity between the dimensions is calculated based on the formula of 6, but only if both values are negative or both values are positive (equivalent to the condition $|C_{ik} - C_{jk}| \leq |C_{ik} + C_{jk}|$). Otherwise, the value of $\frac{|C_{ik}-C_{jk}|}{|C_{ik}+C_{jk}|}$ will be higher than 0, which is unwanted for this similarity measure.

If C_{ik} and C_{jk} are date values, the date difference in days is calculated according to Formula 10. The output of this value is divided by 10.000, and the minimum of this output value and 1 is returned. As most date differences are smaller than 10.000 days, this number was chosen as maximum difference between two dates. A value higher than this number hardly ever occurs, and thus the similarity between dates closer to each other must have a bigger influence on the similarity.

Finally, if none of the conditions above are met, the output for the dimension will be 1.

Entering the values for all dimensions $d \in D$ for both cases, the total $DISSIM$ value can be calculated. This value will always return a dissimilarity value between 0 and 1, and the case with the lowest dissimilarity value is eventually selected and retrieved in this process.

5 Application of AF-SM Case-Based Reasoning

In this section, the details of the combined C-BR model for driver fitness evaluation will be discussed by giving an overview of the model structure, describing the data transformation for the application of AFR and a similarity measure, and providing the details of testing the model's performance.

5.1 Combining Traditional and A Fortiori Case-Based Reasoning

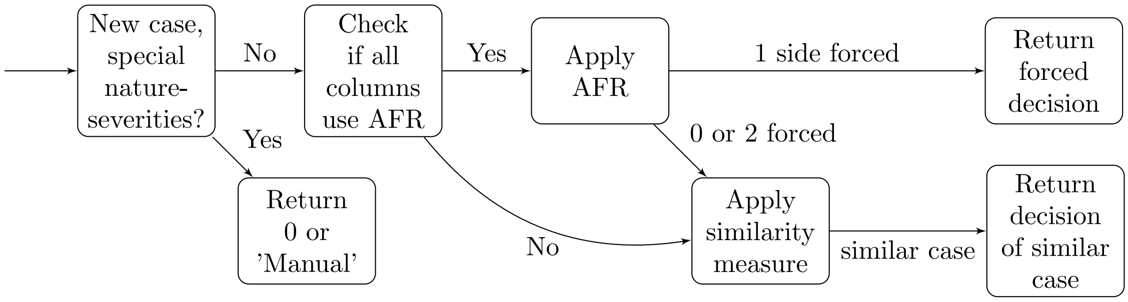

Figure 2: Case retrieval of the combined C-BR algorithm

The new structure of the model, referred to as the combined C-BR model, can be observed in Figure 2 and consists of the following processes. At first, a new problem enters the C-BR model. The initial procedure follows a set of rules to check whether the case is filled with the correct data type per dimension. Furthermore, the case is checked for specific nature-severity combinations. Certain combinations must be decided manually by a medical expert, for which the C-BR models returns it immediately without a decision (see Appendix 9.2). Additionally, certain 'severities' already suggest a shorter validity period of the driver's license. The presence of one of these 'severities' in a case, specified in Appendix 9.3, will always return that an individual is not entirely 'fit to drive'. If no 'manual natures' and 'shorter-validity severities' are present in the case based on these rules, the case is transformed using the case representation 2.1, making the case ready for the application of the AFR process. If none of these nature-severity combinations are present for the case, the case is checked for non-categorical textual dimensions. The presence of one or more of these dimensions will automatically send the case to the application of a similarity measure. In contrast, if all present dimensions are suitable for AFR, the case is sent to the AFR process. This ends the case checking phases of the model.

In the AFR process, the a fortiori algorithm is applied to check whether a decision is forced by another case in the KB by using a fortiori reasoning. This a fortiori algorithm is based on a expansion of the 'AFCBR-algorithm' of [Van Woerkom, 2023], as discussed in Section 4.2.2. However, an additional algorithm was added to retrieve cases that forced a certain decision. This added an explanation in case of a retrieved forced decision.

After checking for a forced decision, there are three possible outcomes: a forced decision is found for one side, no forced decision is found, or forced decisions are found for both decisions. If a forced decision is identified for only one outcome, the corresponding case from the KB that led to this decision is returned along with its outcome. Otherwise, if no forced decision is found or forced decisions are found for both decisions, the retrieval phase continues by looking for one or more most similar cases in the KB with the similarity measure described in Formulas 9 and 8. The most similar case is selected by applying Formula 3, which calculates the minimum dissimilarity value between the new case and all cases in the KB. The case with the lowest dissimilarity value, which has the highest similarity with the new case, is used in the next phase.

All other remaining phases are mostly similar to standard models of C-BR, except for one extra addition: in order to remain as explainable as possible, this C-BR model must clearly indicate whether a decision was forced in the retrieval phase by AFR. This informs the medical expert that the case was not just similar, but a stronger proof for the decision of the new case. If the medical expert does not follow a forced decision and makes a different decision, it creates an inconsistency when saving the current case with the modified solution. This constraint must be avoided.

For a similar case found by the similarity measure, the C-BR model must provide the differences between the new case and the similar case, together with the proposed solution. This statement allows medical experts to carefully examine the differences between cases and determine their own solutions objectively. Besides, cases that forced decisions for the new case are returned as well. This creates the possibility to identify inconsistent cases in the KB, which can be deleted afterwards. The final returned decisions by the C-BR model will act as an advice.

5.2 Data Transformation and Knowledge Base Creation

This section highlights the transformation of the CBR dataset that were made to obtain the necessary components for the case retrieval described C-BR model of Section 5.1.

5.2.1 General CBR Data Transformation

As discussed in Section 3.2.1, every case in the CBR data consists of the present relevant natures and severities in the CBR data. The combination of a nature and a severity with its additional value represents a unique dimension of a case in the CBR database. An example case is given below, in which Table 8 and 9 show the case entry for an example case process x in the CBR database. For the simplicity of explaining the data, an example was chosen which only contained correct values, without any errors.

ID	Nature	Severity	Value	UsedInProcess?
x	DM	DM_BEGINDATUM_GETAL	2018	Yes
x	DM	DM_NIET_INSULINE_HYPO_-	TRUE	Yes
x	VISUS	VISUS_VOD_ZC	0.8	Yes
x	VISUS	VISUS_VOS_ZC	0.8	Yes

Table 8: Nature and Severity data for case x

ID	Decision Date
x	1-1-2023

ID	Category	Decision	ShorterValidity	EUcode
x	B	Geschikt	-	-

Table 9: Decision data for case x

To obtain a sufficient KB, certain filters and data transformations were performed. Obviously, every case process must have a filled 'Decision' column in Table 9. Moreover, all rows that were not used in the decision making process were filtered out of the database (so that 'UsedInProcess? = No' was filtered out). This removed all rows that were irrelevant in the decision making process. Using a sheet with nature-severity information of the CBR, cases that contained invalid nature-severity combinations were filtered out of the KB. Since some nature-severity combinations were no longer used in the decision making process of the CBR, it was reasonable to only use case processes that contained currently used categories of natures and severities. Additionally, all case processes that contained a duplicate nature-severity combination were deleted. For a small number of case IDs rows with duplicate categories were found, although only one entry per category was allowed for every individual. Furthermore, case processes that contained a nature-severity combination that required manual decision making were removed from the dataset. The list of combinations in this category is given in Appendix 9.2.

After all caseIDs with irrelevant information were filtered out, the data was converted into a single row per case ID, as determined in Section 3.2.1. This process yielded a feature-value pair representation as Definition 2.1 for every case. Table 10 illustrates how the data was processed into a case in the KB, after converting the nature-severity combinations to column headers and using their corresponding values as values in those columns. Table 10 only shows the present dimensions for this specific case, while all other (empty) columns are not shown. In the database, these combinations of natures and severities had empty value, representing the absence of the nature-severity combination. This way, all dimensions were easily stored and compared to other cases. Every case had one caseID column, 399 nature-severity combination columns and one column for the decision.

ID	DM__DM_BEG...	DM__DM_NIET...	VISUS..._VOS_ZC	VISUS..._VOD_ZC
x	2018	TRUE	0.8	0.8

Table 10: Case representation of x with all non-empty dimensions

As denoted in the previous section as well, a binary output in this C-BR model was needed, which allowed for the use of a fortiori reasoning. Therefore, the decision column was modified. The decision column assumed one of three values: 'Geschikt' (fit), 'Afzien' (refrain) or 'Ongeschikt' (unfit). The values 'Afzien' and 'Ongeschikt' represented two types of negative decisions, while 'Geschikt' represented a positive decision regarding the fitness to drive. However, a case with 'Geschikt' as its decision could also have a shorter validity, which was represented in the column 'ShorterValidity', or required a modification to the vehicle or a physical aid for the individual, shown in column 'EUcode'. In the case of the presence of a 'shorter validity' value or an EU-code value, an individual was determined not entirely fit to drive. Given that the process had to be converted into a binary classification problem, only individuals who were fully fit to drive (without a shorter validity or EUcode) were given the 'fit to drive' decision. Consequently, only these case processes had a decision of 1 (fit to drive), and all other outcomes had value 0 (not fit to drive). The final binary decisions are summarized in Table 11, and the final case representation of case x is shown in Table 12.

Decision in system	ShorterValidity	EUcode	Binary Decision (Label)
Geschikt	-	-	1
Geschikt	nr_of_years	-	0
Geschikt	-	*code*	0
Geschikt	nr_of_years	*code*	0
Afzien	-	-	0
Ongeschikt	-	-	0

Table 11: Decisions in CBR data and their binary decision in the C-BR model

ID	DM__DM_BEG...	DM__DM_NIET...	VISUS..._VOS_ZC	VISUS..._VOD_ZC	Decision
x	2018	TRUE	0.8	0.8	1

Table 12: Final case representation for x without empty values

The final transformation of the dataset deleted all duplicate rows from the dataset. Since many cases contained exactly dimensions and values, removing duplicate rows provided an increase in efficiency and speed of the algorithm.

The final KB with data stored according to Definition 2.1, appeared in the form of Table 13, where v_d was empty for every nature-severity combinations (dimension) d that was absent, and v_d was filled if the nature-severity combination was present in the CBR database. The variables $n_1\_\_s1$, $n_2\_\_s_2$, etc. represent the nature-severity combinations, which are the dimensions, and l stands for the number of dimensions.

ID	$n_1\_\_s1$	$n_2\_\_s2$	...	$n_l\_\_s_l$	Decision
x	v_1	v_2	...	v_l	0 or 1

Table 13: Case representation

The transformation of the data into the case representation of Table 13 provided a sufficient data structure for the application of a similarity measure. This KB consisted of 30.584 unique caseIDs, from which 16.024 were decided for outcome 1 and 14.560 were given decision 0.

5.2.2 AFR for CBR data

Even though the data was converted to fulfill the case representation definition of Definition 2.1 that made it sufficient for the use of similarity measures, many dimensions of the case were insufficient for the use of AFR as they contained empty values, as denoted in Section 2.3. Therefore, an additional KB had to be developed solely for the execution of AFR, which meant that all unusable dimensions, and cases that contained such dimensions, were deleted from the KB. This removed 276 dimensions from the KB, leaving 123 relevant dimensions that were suitable dimensions for AFR. Additionally, 14.741 cases were removed from the initial KB from Section 5.2.1 because of the presence of a textual nature-severity dimension. Consequently, the final KB for AFR consisted of 15.843 cases, which all consisted of 123 dimensions that allowed for AFR.

Unfortunately, this AFR KB contained many empty cells, even though AFR is unable to work with these empty dimensions due to the preference relations. However, a strategy was devised to fill the empty cells with the most favourable value. Based on the knowledge of the possible dimension values for every dimension available for AFR, empty dimensions were filled with the value with the highest preference for value 1 in the data, as mentioned in Section 4.2.1. Based on the fact that a present feature indicated a deviation of an individual's health, an empty value could be considered to have a preference for the output 'fit to drive'. Thus, an empty value simply represented the best possible value for that dimension. Applying the knowledge of the medical experts, all empty cells were filled with the most preferable dimension value for label 1. This generated a filled KB for the application of AFR.

Additionally, the preference relations for all nature-severity combinations were created. First, the model of Van Woerkom [Van Woerkom, 2023] computed the preference relations automatically by calculating Pearson's correlation for every dimension value. This test calculates the extent to which a value of a feature corresponds to a given decision. If value v of dimension d occurs more often in combination with a positive decision than value w occurs with a positive decision, v prefers the positive decision more relative to w. Hence, Pearson's correlation gives an indication of the preference relations. The orders of every preference relation were checked manually by an expert of the CBR and adjusted in case of incorrect preference orderings.

5.2.3 Similarity Measure for CBR data

For the combined C-BR model and the traditional C-BR model, the developed similarity measure of Section 4.3 was employed. Furthermore, the constructed KB described in Section 5.2.1 was applied, using 399 dimensions per case, including all textual dimensions. The KB for the application of the similarity measure consisted of 30.584 unique caseIDs.

5.3 Experimental Setup

In order to validate that the combined approach outperformed approaches solely applying a similarity measure or AFR, the performance of the combined C-BR model was compared with these C-BR structures. Speaking of the 'performance', this denotes the ability to recreate human decision making for the driver fitness evaluation process. The higher the performance scores of the model, the greater the ability to recreate human decision making in this area. By recreating human decisions, the C-BR model can apply the domain knowledge of the CBR and its ability to reason consistently without human failure, to serve as an advice that reduces the decision making time of the medical expert. Additionally, this would allow for more accurate and consistent decisions in the future, as this model never refrains from its consistent reasoning strategy.

This procedure aimed to evaluate the performance of the new C-BR approach that incorporates both a fortiori reasoning and a similarity measure for the driver fitness evaluation. Its performance was compared against three other C-BR methodologies that are discussed below. The comparison of these approaches is centered around their accuracy in making decisions for cases within a specified domain, thereby establishing the superiority of one method over the others in terms of reliability of their decisions.

5.3.1 Model Implementation Details

Four different C-BR models were be tested for driver classification performance scores: one C-BR model that applied a similarity measure (traditional C-BR model), two a fortiori models (positive and negative AF C-BR) and the newly created combination of traditional and a fortiori C-BR (combined C-BR) were be compared. For every model, a different case retrieval algorithm was developed.

Every model initiated by looking for nature and severity combinations in a case that indicated the necessity of manual evaluation of the driver fitness process. Certain 'natures', which can be found in Appendix 9.2, had to be evaluated manually because of specific rules of the CBR. These cases were not taken into account in the training and testing process of this investigation, to prevent incorrect decision making of future cases.

Additionally, cases sometimes contained certain severities that indicated that an individual should have obtained a shorter validity period of their license. For these cases, every model can automatically returned decision 0, because these severities indicated that the individual was not completely fit to drive. These severities are discussed in Appendix 9.3. After checking for these two exceptions, the models all evaluated the case differently. Further details of the retrieval stages of every model are discussed below.

Model 1: Traditional C-BR (with similarity measure): this approach involved C-BR with the application of a traditional similarity measure to identify the most relevant past case to a new case. The decision of the most similar past case were used to decide for the new case. In case of multiple cases with similar lowest similarity scores, only the first case was returned for the purposes of computational efficiency.

Model 2: Negative AF C-BR: this approach utilized a fortiori reasoning to infer the decision of a new case. Whenever a decision was forced by another case in the KB for one solution, this case was returned together with the side of the forced decision. If no forced decision was found, or if a decision was forced for both decisions ('unfit to drive' and 'fit to drive'), this model automatically generates the negative decision, which is 0 ('unfit to drive'). Cases containing insufficient cases for AFR were automatically given decision 0 as well.

Model 3: Positive AF C-BR: this third model was similar to the Negative AF C-BR model. However, when no forced decision was found, or when forced decisions were found for both decisions, the model returned the positive decision 1 ('fit to drive'). All non-forced cases were decided to be 'fit to drive' by this model. Cases containing insufficient cases for AFR were automatically given decision 1 as well.

Model 4: Combined C-BR: this final and new approach was already discussed, and combined both a fortiori reasoning and a similarity measure to enhance the accuracy of case labeling. Details of this model can be found in Figure 2.

5.3.2 Procedure and Metrics

All four C-BR approaches were implemented and tested independently. As the KB of the CBR data contained 30.584 cases, the test set contained 7.646 cases, which applied the traditional 80/20 split for training and testing model in AI (80% for the creation of the KB, 20% for testing). Each approach was trained and tested on the exact same training and testing sets to obtain comparable results.

After testing every model on similar testing sets using similar training sets, different results were obtained per model. The combinations of the model's decision versus the actual decision per case can could have four different values: a C-BR decision of 1 with an actual decision of 1 (TP), a C-BR decision of 0 with an actual decision of 1 (FN), a C-BR decision of 1 with an actual decision of 0 (FP) and a C-BR decision of 0 with an actual decision of 0 (TN). Table 14 summarizes these outcomes.

Actual / Predicted	Fit to drive (1)	Unfit to drive (0)
Fit to drive (1)	TP	FP
Unfit to drive (0)	FN	TN

Table 14: Different classification outcomes

Based on these variables, the performance of each model was determined. The most common and important performance measure is *accuracy*. To measure the accuracy of the models, we simply calculated the percentage of correctly decided cases. The higher a model's accuracy, the better its performance of the overall driver fitness classification process. The formula of accuracy can be found in Equation 11, which divides the correctly predicted cases (TP and TN) by the total number of cases.

The model's *recall*, also called 'true positive rate' was conducted by dividing the number of correctly decided positive cases ('fit to drive') with the total number of actual positive cases. Its formula is given in Equation 12. A high recall means that the model is able to predict 'fit to drive' when an individual should actually be decided 'fit to drive'.

The *precision*, called 'positive predictive value, was calculated by dividing the number of correctly decided positive cases (TP) by the total number of positively predicted cases (TP + FP). A high precision indicates that the model hardly ever predicts unfit drivers to be fit. The formula is given in Equation 13.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

A detailed analysis was conducted to compare the performance of the four approaches, with a particular focus on the accuracy metric and the precision metric. While the accuracy is most important, the precision score was seen as an important measure in this experiment as well, because we must prevent deciding unfit drivers fit to drive. This is more important than classifying fit drivers unfit (which ratio is calculated with the 'recall' score), since this decision could always be reconsidered later. Therefore, precision is more important than recall in this study.

6 Results

This section presents the comparative analysis of four C-BR models, designated here as the Positive AF C-BR model, the Negative AF C-BR model, the traditional C-BR model, and the combined C-BR model, across three key performance metrics: precision, recall, and accuracy.

The test set consisted of 7646 randomly chosen cases from the entire CBR dataset. From these 7646 cases, 16 cases contained nature-severity combinations that required manual decision making or came with a severity that represented a shorter validity period of the license, which returned the automatic decision of 0. This remains us with 7630 test cases in total. 3730 of the 7630 cases had an actual decision of '1' (fit to drive) and 3900 had decision '0' (unfit to drive). Tables 15, 16, 17 and 18 show the results of every model, where 'Pred' represents the model's decision and 'Actual' the actual decision of a test case.

Pred \ Actual	1	0	
1	3598	2186	**5784**
0	132	1714	**1846**
	3730	**3900**	**7630**

Table 15: Results of positive AFC-BR Model

Pred \ Actual	1	0	
1	1028	52	**1080**
0	2702	3848	**6550**
	3730	**3900**	**7630**

Table 16: Results of negative AFC-BR Model

Pred \ Actual	1	0	
1	3491	387	**3878**
0	239	3513	**3752**
	3730	**3900**	**7630**

Table 17: Results of traditional C-BR model

Pred \ Actual	1	0	
1	3421	360	**3781**
0	309	3540	**3849**
	3730	**3900**	**7630**

Table 18: Results of combined C-BR model

Model	Precision	Recall	Accuracy
Positive AF C-BR	0.622	0.965	0.696
Negative AF C-BR	0.952	0.276	0.639
Traditional C-BR	0.900	0.936	0.918
Combined C-BR	0.905	0.917	0.912

Table 19: Precision, Recall and Accuracy scores per C-BR model

The precision metric calculated the percentage of accurately recognized positive cases among all cases classified as positive by the models. With a precision of 95.2%, the negative AF C-BR model performed best in this comparison, demonstrating its superior capacity to find appropriate scenarios with few false positives. With a precision of 90.5% and 90.0% respectively, the combined C-BR and the traditional C-BR model trailed closely behind, while the Positive AF C-BR model showed lower precision at 62.2%.

Recall assesses the model's ability to identify all actual positive cases within the dataset. The Positive AF C-BR model outperformed the others in this metric with a recall rate of 96.5%. The traditional C-BR model had a recall of 93.6%, slightly exceeding the combined C-BR model, which recorded a recall rate of 91.7%. The Negative AF C-BR model scored an extremely low recall score compared to the other three models, only scoring 27.6% in this metric.

Finally, the accuracy represents the proportion of all correct decisions (both positive and negative) made by the models over the total number of cases. The traditional C-BR achieved the highest overall accuracy at 91.8%, followed closely by the combined C-BR model with a 91.2% accuracy score. The Positive and Negative AF C-BR models showed somewhat lower accuracies at 69.6% and 63.9% respectively. Since accuracy is seen as the most important and widely used performance metric in testing AI-models, the traditional C-BR has shown to be the optimal model for this situation compared to the other three models, followed closely by the combined C-BR model. Table 19 summarizes the performance scores per model.

In total, 2.926 of the 7.630 test cases were decided by a forced decision of the a fortiori algorithm in the models that used AFR. The remaining 4704 cases were determined by using the similarity measure in the combined C-BR model.

Table 20 shows the accuracy for every model that used a fortiori reasoning. While all forced decisions obviously have similar accuracy scores, it can be observed that the combined C-BR model outperformed both other AFR models in the other cases.

Model decisions	Accuracy forced decisions	Accuracy unforced decisions	Overall accuracy
Combined C-BR	$2742/2926 = 93.7\%$	$4219/4704 = 89.7\%$	91.2%
Positive AF C-BR	$2742/2926 = 93.7\%$	$2570/4704 = 54.6\%$	69.6%
Negative AF C-BR	$2742/2926 = 93.7\%$	$2134/4704 = 45.4\%$	63.9%

Table 20: Comparison of A Fortiori models

The performance of both models that used a similarity measure can be observed in Table 21. It can be denoted that the traditional C-BR model outperformed the combined C-BR model on cases where a fortiori reasoning was able to make a decision, showing that using the similarity measure was an even better method for predicting these 2926 cases.

Model	Accuracy possible AF decisions	Accuracy similarity decisions	Overall accuracy
Traditional C-BR	$2788/2926 = 95.4\%$	$4219/4704 = 89.7\%$	91.8%
Combined C-BR	$2742/2926 = 93.7\%$	$4219/4704 = 89.7\%$	91.2%

Table 21: Comparison of the C-BR models that use a similarity measure

7 Discussion

The following section will discuss the obtained results, provides the limitations of the model and this study, and denotes further considerations for the CBR.

7.1 Results

In order to have a clear understanding of the term ‘performance’ in this section, we mean ‘the ability to reproduce human decision making for the driver fitness evaluation task by medical experts of the CBR’. A model that ‘performs’ well, is able to recreate human decisions. The ability of the models reproducing human decisions is discussed later in this section.

Looking at the results in Section 6, it can be observed that the traditional C-BR model and the combined C-BR model certainly outperformed the Positive and Negative AF C-BR models. This can be easily explained by the fact that the Positive and Negative AF C-BR models always made an automatic decision when a fortiori reasoning did not generate a forced decision for one decision. Since the unforced cases were quite spread out over both decisions, this means that both the Positive and the Negative AF C-BR model performed poorly. From these results, we can conclude that the traditional C-BR approach and the newly suggested combined C-BR models are much more accurate compared to models that simply decide all cases equally in case of no forced decisions.

A more unexpected outcome is the fact that the traditional C-BR model outperformed the combined C-BR model, even though the combined C-BR model should have only found stronger evidence for certain decisions by using a fortiori reasoning. In order to draw further conclusions from the performances of a fortiori reasoning compared to the similarity measure, we must dive into the details of the model decisions of both approaches. Table 22 shows the results comparing the two methods on a subset of the test set for which the models used another strategy to determine their driver fitness decisions. This subset consists of the 2926 cases that were forced to one side by the combined C-BR model.

Trad. C-BR \ C. C-BR	Correct	Incorrect
Correct	2656	132
Incorrect	86	52

Table 22: Comparing AFR and the similarity measure on forced decisions

Pred \ Actual	1	0	
1	1028	52	**1080**
0	132	1714	**1846**
	1160	**1766**	**2926**

Table 23: Combined C-BR results for cases with a one-sided forced decision

Pred \ Actual	1	0	
1	1099	77	**1176**
0	61	1689	**1750**
	1160	**1766**	**2926**

Table 24: Traditional C-BR results for cases decided by AFR in the combined C-BR model

Both models achieved a high accuracy on this subset of cases, obtaining an accuracy of 93.7% and 95.4% for the combined C-BR model and the traditional C-BR model, respectively. Even though the accuracy is high, it is remarkable that the model that used a similarity measure was able to predict these cases better than the model that used forced decisions. Tables 23 and 24 show the results per model decision and actual decision for the 2.926 cases. Cases with decision 0 were predicted more accurately in the combined C-BR model, only assigning the wrong decision for 52 cases compared to 77 cases of the traditional C-BR model. However, the mistakes for positive decision 1 were twice as much for this combined C-BR model, incorrectly predicting 132 positive cases compared to 61 of the traditional C-BR model. Overall, the traditional C-BR model outperformed the combined C-BR model.

To understand why more forced decisions were incorrect than decisions by using a similarity measure, we must investigate the incorrectly predicted cases by the combined C-BR model in more detail. Example cases are shown below, where only nature-severity combinations are shown that are present for the cases.

TEST CASE:

Case ID	VISUS__VISUS_VODS_MC	VISUS__VISUS_VODS_ZC	VISUS__VISUS_VOD_MC	VISUS__VISUS_VOD_ZC	VISUS__VISUS_VOS_MC	VISUS__VISUS_VOS_ZC	Label
x	0.8	0.5	0.8	0.3	0.5	0.4	1

FORCED BY:

Case ID	VISUS__VISUS_VODS_MC	VISUS__VISUS_VODS_ZC	VISUS__VISUS_VOD_MC	VISUS__VISUS_VOD_ZC	VISUS__VISUS_VOS_MC	VISUS__VISUS_VOS_ZC	Label
y	0.8	0.5	0.8	0.3	0.8	0.4	0

Figure 3: Example case where AFR made an incorrect decision

Figure 3 shows the first example case. In this example, the first row is the test case, where x represents the caseID. The second row shows the case in the KB that led to a forced decision. The columns in the picture show the present nature-severity combinations for both cases, and the 'Label' column represents the actual decision. All 'irrelevant' empty dimensions were left out of the comparison, as they were filled with the most preferable value for decision 1 for both cases. In this case, the actual decision for case x was 1, while the use of AFR forced decision 0. It can be observed that all present dimensions contained equal values for both cases, except for the 'VISUS__VISUS_VOS_MC' dimension. In this example case, all dimensions preferred decision 1 over 0, so a higher value preferred being 'fit to drive' over a lower value. Based on case y, a fortiori reasoning concluded that the decision for case x should have been 0, as all its values are lower than or equal to the values of y. Therefore, the combined C-BR model forced decision 0. This mistake indicates that the KB was not consistent, which means that not every case is evaluated equally by medical experts. Even though almost all VISUS values were equal and one value was worse in case x, the case was still decided 'fit to drive', opposing the suggested decision by case y.

In the AFCBR model of Van Woerkom [Van Woerkom, 2023], a measure was generated to calculate the consistency of a KB. As the example case above was incorrectly decided, the consistency percentage of this KB provides information regarding double-sided forced decisions, or incorrect one-sided forced decisions. The consistency percentage is "the relative frequency of cases in the case base that have their outcome forced for the outcome they did not receive". Results of this measure show that the KB of the CBR only comes with a consistency percentage of 54.5%, which means that only 8.642

of the 15.843 AFR cases were consistent. In order to increase the accuracy of the combined C-BR model, the CBR has to determine which cases must be removed from the current dataset, to prevent incorrect forced decisions.

TEST CASE:

Case ID	VISUS__VISUS_VOD_MC	VISUS__VISUS_VOD_ZC	VISUS__VISUS_VOS_MC	VISUS__VISUS_VOS_ZC	Label
x	0.8	0.12	0.3	0.5	1

FORCED BY

Case ID	BIPOLAIRE_STOORNIS__BIP_STOORN_DATUM_DIAG	Label
y	1992-01-01	0

Figure 4: Another example case where AFR predicted incorrectly

However, human inconsistency was not the only reason for incorrect decisions of the combined C-BR model. The example in Figure 4 shows another reason. In this example we see that case y incorrectly forced decision 0 for case x, even though the nature-severity combination 'BIPOLAIRE_STOORNIS__BIP_STOORN_DATUM_DIAG' was not even present in case x, and not relevant for the decision of case x. This incorrect decision was generated because of mistakes in the initialization of the most preferred values in the model, which was done incorrectly for this nature-severity combination. Every nature-severity combination obtained a preferred when a combination was missing based on the preference relations, denoted in 4.2.1, to be able to use AFR in this case study. However, by this incorrect initialization of the 'BIPOLAIRE_STOORNIS__BIP_STOORN_DATUM_DIAG' dimension, the model reasoned that the default value of case x was worse than the value of y. In addition, the default values for all 'VISUS' nature-severity combinations was set to 1, which meant that all VISUS values for case x were seen worse than case y.

Among the inconsistent evaluation of driver fitness by medical experts of the CBR, which causes inconsistent KB data, and the incorrect initialization of default values for missing nature-severity combinations, there exist multiple other causes for the incorrect forced decisions by the AFR algorithm. Firstly, the test case could have an inconsistent decision compared to other cases in the KB. However, this would result in incorrect decisions of the traditional C-BR model as well. Secondly, the preference relations could have been incorrectly generated. Even though the medical knowledge is implemented in the AFR algorithm, it is still possible that the preference relation of certain nature-severity combinations are stored incorrectly. This would result into incorrectly ordering certain values above others, when an opposing order would be correct, resulting in incorrect forced decisions. Finally, incorrect data storage causes incorrect decisions by both the AFR algorithm and the similarity measure.

All the reasons mentioned above cause the AFR algorithm to incorrectly force certain decisions. For these reasons, the number of double-sided or incorrect forced decisions in this case study is considerably high. Further identification of inconsistent cases and incorrectly generated preference relations could improve the model by deleting these cases from the KB and modifying the preference relations. In order to determine the instances in which preference relations ought to be revisited according to the specific nature-severity combination in question, it is necessary to consider the frequency of occurrence of a given column in cases where the AFR model was incorrectly predicted. Table 25 shows the number of times a column was present in an incorrectly decided case, the total appearances in the AFR-decided cases and the ratio of mistakes. The entire table can be found in Appendix 9.4. The ratio together with the 'number of mistakes' column shows which nature-severity combinations require a reconsideration of the preference relation.

Nature-severity combination	Incorrect forces	Total forces	%
REUMA__REUMA__RA	4	4	1,00
EPILEPTISCHE_AANVAL_LEN)__EPILEP_EENV_LAATSTE_AANV_DAT	1	1	1,00
ANGINA_PECTORIS__ANG_PECT_INSTAB	1	1	1,00
OOGAANDOENING__OOGAANDOEN_NYSTAGMUS_VERW	1	1	1,00
CNS__CNS_IDOPATISCH	1	1	1,00
ORGAANTRANSPLANTATIE__LEVER	1	1	1,00
HARTRITMESTOORNIS__HARTRITMESTOORN_BRADY	2	3	0,67
EPILEPTISCHE_AANVAL_LEN)__ANTI_EPILEP_STOPDATUM	10	19	0,53
CNS__CNS_NF_PERCENTAGE	7	14	0,50
BIPOLAIRE_STOORNIS__BIP_STOORN_DATUM_DIAG	1	2	0,50
REUMA__REUMA_WEKE_DELEN	1	2	0,50
OOGAANDOENING__OOGAAND_ABL_RET_LIST	3	7	0,43
HARTKLEP_AFWIJKING__HARTKLEP_AFW_PROTHESE	9	26	0,35
DM__DM_ERNSTIG_HYPO_1_JAAR	1	3	0,33
CNS__CNS_MDRD_STAD_IV_20_29	1	3	0,33
EPILEPTISCHE_AANVAL_LEN)__EPILEP_SPOR_INT_GD_2_JAAR	3	11	0,27
EPILEPTISCHE_AANVAL_LEN)__EPILEP_ANDERE_LAATSTE_AANV_DAT	8	30	0,27

Table 25: Occurrences of columns in incorrectly predicted cases by a fortiori reasoning compared to the total number of occurrences in the AFR possible cases

On the other hand, the combined C-BR model sometimes applies its reasoning correctly where the similarity measure lacks predictability. Thus, it remains important to identify cases in which the model outperformed the traditional C-BR model. An example case is given below, in which the similarity measure found a similar case with worse VISUS values with decision 0, and AFR found a forced decision for decision 1.

TESTING CASE

Nature__Severity	PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_CATARACT	VISUS__VISUS_VODS_MC	VISUS__VISUS_VODS_ZC	VISUS__VISUS_VOD_MC	VISUS__VISUS_VOD_ZC	VISUS__VISUS_VOS_MC	VISUS__VISUS_VOS_ZC	Label
ZaakProcesnummer								
x	OOG NIET BEKEND	1.5	1.5	1.5	0.8	1.5	1	1

AFR CASE

AardErnstAardErnst	PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_CATARACT	VISUS__VISUS_VODS_MC	VISUS__VISUS_VODS_ZC	VISUS__VISUS_VOD_MC	VISUS__VISUS_VOD_ZC	VISUS__VISUS_VOS_MC	VISUS__VISUS_VOS_ZC	Label
y	OOG NIET BEKEND	1.25	0.8	1.25	0.8	0.8	0.5	1

SIM CASE

AardErnstAardErnst	PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_CATARACT	VISUS__VISUS_VODS_MC	VISUS__VISUS_VODS_ZC	VISUS__VISUS_VOD_MC	VISUS__VISUS_VOD_ZC	VISUS__VISUS_VOS_MC	VISUS__VISUS_VOS_ZC	Label
z	OOG NIET BEKEND	1.5	1.25	1	0.8	1.5	1	0

Figure 5: Example case where AFR made the decision correctly and the similarity measure incorrectly

Now that the cases are discussed where a fortiori could be applied, it remains important to investigate the performance of cases decided by the similarity measure as well. In total, 4.704 test cases were undecided by a fortiori reasoning, which was caused by one of the following three reasons: the case was forced for both decisions by the a fortiori algorithm (1.974 of the 4.704 cases), the case was forced for no decision (73 of the 4.704 cases), or the case contained nature-severity combinations that could not be handled in a fortiori reasoning (2657 of the 4.704 cases). All cases from these three categories were decided by applying the similarity measure. Results per category are shown in Tables 26, 27 and 28.

Pred \ Actual	1	0	
1	1046	59	**1105**
0	51	818	**869**
	1097	**877**	**1974**

Table 26: Combined C-BR similarity results for cases that achieved double AFR forced decisions

Pred \ Actual	1	0	
1	8	9	**17**
0	3	53	**56**
	11	**62**	**73**

Table 27: Combined C-BR similarity results for cases that achieved no AFR forced decisions

Pred \ Actual	1	0	
1	1339	240	**1579**
0	123	955	**1078**
	1462	**1195**	**2657**

Table 28: Combined C-BR results for cases that contained columns not suitable for AFR

Model	Accuracy
Forced two sides	94.4%
Forced no sides	83.6%
Unsuitable for AFR	86.3%

Table 29: Accuracy per category

First, the unsuitable cases for a fortiori reasoning are discussed, which are found in Table 28. These cases contained nature-severity values that could not be handled by a fortiori reasoning algorithms, such as textual explanations of medical deviations. It can be observed that the majority of cases that were undecided by a fortiori reasoning in the combined C-BR model contained these textual nature-severity combinations, since 2.657 of the 4.704 undecided AFR fell in this category. For these 2.657 cases, 2.294 cases were decided correctly, achieving an accuracy percentage of 86.3%. Even though the similarity measure did not incorporate any understanding of natural language, it can be concluded that checking for equality in these columns contributed to a quite high accuracy percentage.

Diving further into the cases that were technically possible to solve with a fortiori reasoning in terms of present nature-severity combinations, this leaves us with 2.047 cases of the test set. These cases were solved by a similarity measure as well, since not exactly one decision was forced. Only 73 of these cases did not obtain a forced decision at all, while 1957 cases had forced decisions for both sides. Results for these cases can be found in Table 26 and 27. For the 73 unforced cases, 62 had an actual decision of 'unfit to drive' and only 11 were actually decided 'fit to drive'. From this, we can conclude that in case of an actual positively decided case, the model was more likely to find a forced decision than for a negatively decided case.

Another interesting outcome can be found in the possible a fortiori cases from Tables 23 26 and 27. In total, 4.973 cases contained nature-severity combinations that could all be handled by AFR. From the 4973 cases in which the a fortiori reasoning algorithm was applied, 1.957 cases obtained a forced decision for both sides, 2.926 cases for only one side and 73 for none of the sides. A double-sided forced decision for 39.4% of the a fortiori test cases means that the KB for a fortiori reasoning is far from being consistent. This can be concluded as well when looking at the percentage of a fortiori decisions, which is 6.6%. Therefore, it remains important to remove incorrect and inconsistent cases from the KB, which makes decisions by the model in the future possibly more accurate.

As denoted above, a review of the data revealed that 39.4% of all AFR-applied cases were forced for both decisions. This prompts the question of why so many cases were forced for both decisions. Figure 6 shows an example case, which can be found in the top rows of the figure. All rows again solely contain the present nature-severity combinations of the individuals. All 'VISUS' columns prefer decision 1, which again implies that higher values are more likely to obtain a decision of 1. The values of 'PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_CATARACT' are equal in every case, all being 'OOG NIET BEKEND' (meaning 'particular eye is unknown'). While case y forces decision 0 for case x because the values of y are better than or equal to the values of x, case z forces the opposite decision 1 having values worse than or equal to case x. This is a great example showing the inconsistencies that appear in the dataset of the CBR. These inconsistencies render the AFR model incapable of forcing decisions in many cases. A manual evaluation of all double-sided forced decisions together with the removal of incorrect decisions could increase the accuracy of the AFR decisions.

CaseID	PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_CATARACT	VISUS__VISUS_VODS_MC	VISUS__VISUS_VODS_ZC	VISUS__VISUS_VOD_MC	VISUS__VISUS_VOD_ZC	VISUS__VISUS_VOS_MC	VISUS__VISUS_VOS_ZC	Label
x	OOG NIET BEKEND	2	1	1	1	1.5	0.5	1

AFR for Label 0

CaseID	PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_CATARACT	VISUS__VISUS_VODS_MC	VISUS__VISUS_VODS_ZC	VISUS__VISUS_VOD_MC	VISUS__VISUS_VOD_ZC	VISUS__VISUS_VOS_MC	VISUS__VISUS_VOS_ZC	Label
y	OOG NIET BEKEND	2	1.25	1.5	1.25	1.5	0.8	0

AFR for Label 1

CaseID	PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_CATARACT	VISUS__VISUS_VODS_MC	VISUS__VISUS_VODS_ZC	VISUS__VISUS_VOD_MC	VISUS__VISUS_VOD_ZC	VISUS__VISUS_VOS_MC	VISUS__VISUS_VOS_ZC	Label
z	OOG NIET BEKEND	1	0.63	1	0.63	1	0.4	1

Figure 6: An example of case with double forced decisions

Overall, the combined C-BR model proved to be almost as accurate in reproducing human decisions as the traditional C-BR model that only used a similarity measure. However, it has not proven yet to perform better than simply using a similarity measure to find an explanation for the AI-model's decision. Future work could investigate whether the combination of a fortiori reasoning and similarity measures in C-BR models might improve the accuracy in other domains. The CBR database contained complex data, which might have increased the difficulty of preprocessing the data. Therefore, the domain of this case study may not be the most suitable situation of testing the combination of the two case retrieval approaches.

Additionally, a reconsideration of decisions for cases that obtained double-sided forced decisions might significantly increase the accuracy of the combined C-BR model. This would improve the usefulness of the combined C-BR model for the CBR, as the model will be able to make even more accurate decisions similar to decisions by medical experts of the CBR, making the reasoning of the model more consistent.

Finally, even though the combined C-BR model is quite accurate, this does not imply that the developed C-BR model might be sufficient for advising medical experts of the C-BR. Being able to reproduce human decisions in this domain does not automatically infer that decisions will be made more consistent in the future. This might only happen if inconsistent cases are recognized by medical experts of the C-BR, and if these are removed from the KB. This will result create a more consistent KB, which could increase the accuracy of the AFR decisions. Thus, the accuracy of the advised decisions of the model could improve with the removal of inconsistent cases from the KB.

7.2 Limitations

Even though the combined C-BR model already reproduces human decisions by medical experts in 91.2% of the times in this case study, the model comes with further limitations. These limitations will be discussed below.

Firstly, the model requires detailed knowledge of the data itself to insert knowledge into the KB of the C-BR model. Preference relations for the a fortiori algorithm in the model were constructed automatically by the algorithm of Van Woerkom [Van Woerkom, 2023] and adjusted manually by determining an order for the values of every dimension of the cases. The automatic assignment of preference relations to every dimension showed that the generated relations did not always represent the true preference of every dimension. Thus, the automatic assignment has its limitations in terms of accurate preference relation generation. However, the automatic assignment did reduce the development time compared to manual preference creation.

Additionally, an addition of adjustment of nature-severity combinations would change the representation of all cases, as these have to be adjusted as well. Furthermore, non-linear variables cannot be captured in this version of the a fortiori algorithm, which means that non-linear numerical dimensions have to be preprocessed to obtain column values that can be handled by the a fortiori algorithm.

Moreover, the KB and the test set only contained a small number of cases compared to the entire dataset of the CBR. Due to limited time and resources, the model could not be trained and tested on more cases, which would have given more information about the performance of the model. However, this case study already shows that the model is able to perform well on the CBR data.

Finally, the lack of interpretability of textual dimension values caused the model to not perform optimally on cases containing such dimensions. Examples of these textual features are fields where medical specialists describe certain deviations of individuals, which are simply impossible to 'understand' for this C-BR model. In the current combined C-BR model, textual values are only considered similar if the entire text is exactly similar to the text of another case. Partial similarity is not implemented in this C-BR model, since the similarity measure should include NLP to understand textual values of certain natures and severities. Because of the efficiency and speed of the current model, these textual descriptions are left out of consideration in the similarity calculations of this C-BR model, leaving space for further improvement.

An additional limitation of both the traditional C-BR model and the combined C-BR model can be found in the similarity measure process. In case of two or more cases that have an equal best similarity score, only the first case is returned. This is because the computational speed dropped significantly if multiple cases had to be returned in this situation. However, this could have resulted in a sub-optimal decision making process of the C-BR model, as cases with equal similarity scores should have all been taken into account in the final decision.

7.3 Considerations for the CBR

Even though the combined C-BR model has shown to be able to reproduce human decisions in 91.2% of the cases, the model is not yet sufficient to serve as an advisory tool. Certain aspects of the model must be taken into consideration before applying this model.

Firstly, as already discussed before, every decision by the C-BR model is simply an advice on the fitness to drive of an individual. This decision may never be directly adopted as the final decision under any circumstances, because of the possibility of the 'control problem' denoted in Section 2.2. Relying solely on the knowledge and reasoning of the C-BR model could cause inaccurate decision making. While this impacts the decision of one individual, the adaptation of an incorrect decision in the KB of the C-BR model will additionally negatively influence the decision making of future cases. Thus, manual checking of the decisions by medical experts is a must.

Secondly, the addition and removal of features of cases must be done carefully. The adaptation of dimensions provides for adjustments in the storage of cases in the KB of the C-BR model and in the development of preference relations for this dimension. A default value must be inserted in the KB, a preference relation has to be stored and all previous and new cases must contain a new column for this added nature-severity combination.

Finally, the C-BR model was not developed to improve the decision making in driver fitness evaluation. Case-based reasoning applies previously evaluated cases, which means that the model will initially not be able to make better decisions than humans. However, if the medical experts can remove incorrectly decided cases from the KB, and decide which of the inconsistent cases is still a correct decision, the accuracy of the C-BR model will improve. If the inconsistent cases are removed from the KB and the preference relations are implemented so that it represents the knowledge of the CBR domain, the model might possibly decide more consistently than humans in the future: as intuitive or emotional responses do not play a significant role in C-BR models decision-making, human reasoning might more sensitive to these weaknesses. However, because of the reliability of C-BR on manually decided cases, it remains important for medical experts to decide consistently. Otherwise, cases with similar dimension values could come with different final decisions, which would make the driver fitness evaluation process even less consistent, causing the model to perform even worse.

8 Conclusion

In this case study at the CBR, it was investigated how the combination of a fortiori reasoning and similarity measure in case-based reasoning contributes to the ability of C-BR to recreate human decisions compared to more standard C-BR models. In order to answer this question, the following sub-questions were answered.

Q1. What are the advantages of combining a fortiori reasoning and a similarity measure in C-BR?

Combining a fortiori reasoning and similarity measures in C-BR could improve the accuracy of decisions generated by the model and make its decisions more justifiable than models solely utilizing a similarity measure. Since AFR only decides for cases that consist of better evidence for a previous decision, which are called forced decisions, this suggests that AFR returns correct decisions if previous decisions are decided correctly. Besides, these decisions could be more justifiable, as the decision is only made for cases that are better evidence for a previously decided case according to the KB. A similarity measure simply returns the decision of the case that is closest to the new case, which might not even be similar to the new case. Therefore, AFR might improve the accuracy and justifiability of a C-BR model. However, the ability of AFR to create correct forced decisions is dependent on the consistency of the KB.

Additionally, the combination of a fortiori reasoning and a similarity measure in C-BR will be able to decide for more new cases than C-BR solely utilizing a fortiori reasoning. This is because not every case will receive exactly one forced decision for one of the possible sides. If the case contains dimensions that are insufficient for the use of a fortiori reasoning, the case will be left undecided. Similarly, if decisions are forced for no side or for both sides, a fortiori reasoning will not return a decision. For these undecided cases, a similarity measure still provides a solution by looking for the most similar case in the KB and returning its decision.

The integration of AFR in C-BR, followed by the application of SM for undecided cases, will result in the creation of a combined C-BR model that incorporates the advantages of both approaches.

Q2. How can a fortiori reasoning be utilized in the domain of this study?

In the domain of the CBR, a fortiori reasoning can be applied by transforming the dataset of the CBR to obtain feature-value pairs. Since the data is already structured likewise, only removing errors and unwanted values from the dataset will generate a knowledge base that is sufficient for the use of a fortiori reasoning. Because of the presence of many empty values in the data, these dimension values must be filled with the most preferred value for that dimension to obtain a positive fitness to drive decision, since the absence of a value denotes the absence of a health deviation. Thus, filling empty values with the 'best' value for a positive decision removes any negative influence of that dimension towards the final decision. Finally, by developing preference relations for every suitable dimension, a fortiori reasoning can be applied on the knowledge base. The AFCBR (a fortiori case-based reasoning) model of Van Woerkom [Van Woerkom, 2023] serves as a practical and quick implementation of a fortiori reasoning in Python, which was modified for the current case study.

Q3. How can a similarity measure from the C-BR literature be modified to compare cases in the domain of this study?

Because the CBR data consists of many empty values and other data types, an existing similarity measure was adjusted for the use of cases from the CBR. The initial similarity measure that accounted for other data types by Castro et al. [Castro et al., 2009] was modified, as this similarity measure is widely used in other C-BR studies. As the impact of every dimension to the output was unknown, the weight of every dimension was set to 1. Besides, the dissimilarity between each dimension value was adjusted to allow for more data types and empty values. This measure was given in Section 4.3.

Q4. How is the performance of the combined C-BR model measured to compare the model to C-BR models solely utilizing a similarity measure or a fortiori reasoning?

The current decision making process of the CBR relies on the judgment of medical experts of the CBR. These experts apply their knowledge and expertise to determine the fitness to drive of individuals. For every individual, the present health deviations are taken into account, together with their values. To measure the performance of each C-BR model on this driver fitness assessment, the performance of each model was be tested on cases already evaluated by medical experts of the CBR. By testing on these cases, the ability of each C-BR model to reproduce human decisions was assessed. By comparing the models' decisions with the decisions from medical experts of the CBR, the accuracy of these C-BR decisions were measured. This allowed us to determine the extent to which each model was suitable for reproducing human decisions. In addition, the precision and recall scores were measured.

Even though the data of the CBR consists of cases that were manually decided by medical experts, these cases could still be a sufficient basis for the KB of a C-BR model. While human inconsistencies appear in these manual decisions, these cases can be deleted from the KB in the future, through the recognition of incorrect decisions by the medical experts themselves. When the C-BR model creates an incorrect advisory decision based on one or more cases from the KB, these incorrect cases can be deleted from KB of the model. This will make the decision-making process of the C-BR model more consistent and increase the usability of C-BR in this domain.

The previous questions were answered in this thesis to be able to develop and test the combined C-BR model. Results have shown that the combined C-BR model did not outperform a traditional C-BR approach that only applied a similarity measure to predict new cases, based on its accuracy, precision and recall score. On the other hand, there is room for improvement, and the current accuracy only deviated 0.6% from the 91.8% of the traditional C-BR model that solely used a similarity measure. Nevertheless, these results suggest that one of the developed models might be a sufficient tool for assisting medical experts of the C-BR in the future, to later obtain more consistent decision making and possibly reduce the time used per case. An overall accuracy of 91.2% concludes that the C-BR model will predict new cases correctly most of the time. From this, and the fact that a C-BR model is able to decide much quicker than humans, it can be concluded that this model might serve as a useful tool to advise medical experts of the CBR in the future. The model could already generate an advice when all medical deviations are determined, and provide evidence for its decision based on past cases. Nevertheless, there is always room for improvement in this area, as the traditional C-BR model performed even better. Thus, the accuracy of the model could be improved, which would make the model even better at assisting medical experts of the CBR.

Future work in the CBR domain could investigate if the removal of inconsistent cases in the C-BR model could improve the performance of the combined C-BR model, as this will allow for fewer double-sided forced decisions. Furthermore, the preference relations might not capture the preference of every dimension correctly. Adjustments of these preference relations and the KB might cause the model to perform better. These modifications might potentially increase the performance of a fortiori reasoning in the combined C-BR model for the CBR. Even though the model is not sufficient yet to serve as an advisory tool for the CBR, removing inconsistent cases from the KB and reconsidering the preference relations might possibly make the reasoning process of C-BR more consistent than human decision making in the future: as intuitive and emotional responses do not play a role in the decision making process of C-BR, these models will not be subjective to human inconsistencies if there is no inconsistent data in the KB. Thus, the AFR process of the combined C-BR model will only improve if inconsistencies are deleted from the C-BR model.

Taking a broader scope, further studies could investigate the application of this proposed combined C-BR model in other domains. Even though the performance of the traditional C-BR model was slightly higher in this domain, it can be concluded that the combination of AFR and a similarity measure in C-BR provides a hopeful combination of the legal C-BR approach and the more standard C-BR approach. The complexity of the dataset of the CBR made it challenging to transform the data in a way that would enable C-BR, let alone to test the newly generated model. Therefore, it is of interest to test the performance of a combined C-BR model in a domain where the data is less complex, but especially more complete and simpler. In a database where all features are filled and fewer features are considered in case comparison, there is no need to fill in empty values using manually created rules. Moreover, the preference relations per feature in a fortiori reasoning can be defined and adjusted more easily, since fewer features are taken into account. Testing the combined C-BR approach in other domains might prove its potential

Another interesting research is to develop a similarity measure that applies the preference relations in its similarity calculation. For example, by calculating the number of dimensions per case that prevented a forced decision. In this study, the similarity measure did not apply the preference relations of AFR. The addition of these preference relations to similarity measures would combine the two approaches even more. By counting the ratio of preferred dimension values per case, these preference relationships could be applied in the similarity measures as well, providing more unity between the methods and possibly a new useful similarity measure.

This thesis contributed to the scientific field of AI by providing a new case-based reasoning approach that combined the advantages of legal C-BR and more traditional C-BR using a similarity measure. Even though this model does not yet outperform previously suggested C-BR models, this study gives hope for further combined C-BR studies.

References

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

[Aleven, 1997] Aleven, V. A. (1997). *Teaching case-based argumentation through a model and examples*. Citeseer.

[Balasubramaniam et al., 2022] Balasubramaniam, N., Kauppinen, M., Hiekkanen, K., and Kujala, S. (2022). Transparency and explainability of AI systems: ethical guidelines in practice. In *International Working Conference on Requirements Engineering: Foundation for Software Quality*, pages 3–18. Springer.

[Batista and Monard, 2003] Batista, G. E. and Monard, M. C. (2003). An analysis of four missing data treatment methods for supervised learning. *Applied artificial intelligence*, 17(5-6):519–533.

[Bergmann et al., 2005] Bergmann, R., Kolodner, J., and Plaza, E. (2005). Representation in case-based reasoning. *The Knowledge Engineering Review*, 20(3):209–213.

[Bjørner et al., 2019] Bjørner, N., de Moura, L., Nachmanson, L., and Wintersteiger, C. M. (2019). Programming z3. *Engineering Trustworthy Software Systems: 4th International School, SETSS 2018, Chongqing, China, April 7–12, 2018, Tutorial Lectures 4*, pages 148–201.

[Blanco Valencia et al., 2018] Blanco Valencia, X., Bastidas Torres, D., Piñeros Rodriguez, C., Peluffo-Ordóñez, D. H., Becerra, M. A., and Castro-Ospina, A. E. (2018). Case-based reasoning systems for medical applications with improved adaptation and recovery stages. In *Bioinformatics and Biomedical Engineering: 6th International Work-Conference, IWBBIO 2018, Granada, Spain, April 25–27, 2018, Proceedings, Part I 6*, pages 26–38. Springer.

[Buchanan, 2005] Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. *Ai Magazine*, 26(4):53–53.

[Canavotto and Horty, 2022] Canavotto, I. and Horty, J. (2022). Piecemeal knowledge acquisition for computational normative reasoning. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, pages 171–180.

[Castro et al., 2009] Castro, J. L., Navarro, M., Sánchez, J. M., and Zurita, J. M. (2009). Loss and gain functions for cbr retrieval. *Information Sciences*, 179(11):1738–1750.

[Chattopadhyay et al., 2013] Chattopadhyay, S., Banerjee, S., Rabhi, F. A., and Acharya, U. R. (2013). A case-based reasoning system for complex medical diagnosis. *Expert Systems*, 30(1):12–20.

[Cunningham, 2008] Cunningham, P. (2008). A taxonomy of similarity mechanisms for case-based reasoning. *IEEE Transactions on Knowledge and Data Engineering*, 21(11):1532–1543.

[Cutillo et al., 2020] Cutillo, C. M., Sharma, K. R., Foschini, L., Kundu, S., Mackintosh, M., Mandl, K. D., Beck, T., Collier, E., Colvis, C., Gersing, K., et al. (2020). Machine intelligence in healthcare—perspectives on trustworthiness, explainability, usability, and transparency. *NPJ Digital Medicine*, 3(1):47.

[De Martino et al., 2006] De Martino, B., Kumaran, D., Seymour, B., and Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, 313(5787):684–687.

[El Naqa and Murphy, 2015] El Naqa, I. and Murphy, M. J. (2015). *What is machine learning?* Springer.

[El-Sappagh and Elmogy, 2015] El-Sappagh, S. H. and Elmogy, M. (2015). Case based reasoning: Case representation methodologies. *International Journal of Advanced Computer Science and Applications*, 6(11):192–208.

[Finnie and Sun, 2002] Finnie, G. and Sun, Z. (2002). Similarity and metrics in case-based reasoning. *International journal of intelligent systems*, 17(3):273–287.

[Finnie and Sun, 2003] Finnie, G. and Sun, Z. (2003). R5 model for case-based reasoning. *Knowledge-Based Systems*, 16(1):59–65.

[Gerlings et al., 2020] Gerlings, J., Shollo, A., and Constantiou, I. (2020). Reviewing the need for explainable artificial intelligence (XAI). *arXiv preprint arXiv:2012.01007*.

[Gilpin et al., 2018] Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., and Kagal, L. (2018). Explaining explanations: An overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*, pages 80–89. IEEE.

[Goldberg et al., 2008] Goldberg, S. I., Niemierko, A., and Turchin, A. (2008). Analysis of data errors in clinical research databases. In *AMIA annual symposium proceedings*, volume 2008, page 242. American Medical Informatics Association.

[Heindl et al., 1997] Heindl, B., Schmidt, R., Schmid, G., Haller, M., Pfaller, P., Gierl, L., and Pollwein, B. (1997). A case-based consiliarius for therapy recommendation (icons): computer-based advice for calculated antibiotic therapy in intensive care medicine. *Computer Methods and Programs in Biomedicine*, 52(2):117–127.

[Holzinger et al., 2017] Holzinger, A., Biemann, C., Pattichis, C. S., and Kell, D. B. (2017). What do we need to build explainable AI systems for the medical domain? *arXiv preprint arXiv:1712.09923*.

[Horty, 2011] Horty, J. F. (2011). Rules and reasons in the theory of precedent. *Legal Theory*, 17(1):1–33.

[Horty and Bench-Capon, 2012] Horty, J. F. and Bench-Capon, T. J. (2012). A factor-based definition of precedential constraint. *Artificial intelligence and Law*, 20:181–214.

[Janiesch et al., 2021] Janiesch, C., Zschech, P., and Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3):685–695.

[Jiang et al., 2017] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., and Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, 2(4).

[Kang et al., 2013] Kang, Y.-B., Krishnaswamy, S., and Zaslavsky, A. (2013). A retrieval strategy for case-based reasoning using similarity and association knowledge. *IEEE transactions on cybernetics*, 44(4):473–487.

[Kolodner, 2014] Kolodner, J. (2014). *Case-based reasoning*. Morgan Kaufmann.

[Kolodner, 1992] Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial intelligence review*, 6(1):3–34.

[Kulikowski, 1988] Kulikowski, C. A. (1988). Artificial intelligence in medical consultation systems: A review. *IEEE Engineering in Medicine and Biology Magazine*, 7(2):34–39.

[Lance and Williams, 1966] Lance, G. N. and Williams, W. T. (1966). Computer programs for hierarchical polythetic classification ("similarity analyses"). *The Computer Journal*, 9(1):60–64.

[Liao et al., 1998] Liao, T. W., Zhang, Z., and Mount, C. R. (1998). Similarity measures for retrieval in case-based reasoning systems. *Applied Artificial Intelligence*, 12(4):267–288.

[London, 2019] London, A. J. (2019). Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Center Report*, 49(1):15–21.

[Miller, 1986] Miller, P. L. (1986). The evaluation of artificial intelligence systems in medicine. *Computer methods and programs in biomedicine*, 22(1):3–11.

[Morrow and Sormani, 2020] Morrow, J. M. and Sormani, M. P. (2020). Machine learning outperforms human experts in mri pattern analysis of muscular dystrophies. *Neurology*, 94(10):421–422.

[Núnez et al., 2004] Núnez, H., Sanchez-Marre, M., Cortés, U., Comas, J., Martínez, M., Rodríguez-Roda, I., and Poch, M. (2004). A comparative study on the use of similarity measures in case-based reasoning to improve the classification of environmental system situations. *Environmental Modelling & Software*, 19(9):809–819.

[Odekerken and Bex, 2020] Odekerken, D. and Bex, F. (2020). Towards transparent human-in-the-loop classification of fraudulent web shops. In Villata, S., Harasta, J., and Kremen, P., editors, *Legal Knowledge and Information Systems. JURIX 2020: The Thirty-Third Annual Conference*, volume 334, pages 239–242. IOS Press.

[Park et al., 2004] Park, J. H., Im, K. H., Shin, C.-K., and Park, S. C. (2004). Mbnr: case-based reasoning with local feature weighting by neural network. *Applied Intelligence*, 21:265–276.

[Peters et al., 2023] Peters, J. G., Bex, F. J., and Prakken, H. (2023). Model-and data-agnostic justifications with a fortiori case-based argumentation. In *19th International Conference on Artificial Intelligence and Law*, pages 207–216. ACM Press.

[Poesio, 1995] Poesio, M. (1995). Semantic ambiguity and perceived ambiguity. *arXiv preprint cmp-lg/9505034*.

[Richter and Aamodt, 2005] Richter, M. M. and Aamodt, A. (2005). Case-based reasoning foundations. *The Knowledge Engineering Review*, 20(3):203–207.

[Richter and Weber, 2016] Richter, M. M. and Weber, R. O. (2016). *Case-based reasoning*. Springer.

[Samek and Müller, 2019] Samek, W. and Müller, K.-R. (2019). Towards explainable artificial intelligence. *Explainable AI: interpreting, explaining and visualizing deep learning*, pages 5–22.

[Schoenborn et al., 2021] Schoenborn, J. M., Weber, R. O., Aha, D. W., Cassens, J., and Althoff, K.-D. (2021). Explainable case-based reasoning: A survey. In *AAAI-21 Workshop Proceedings*.

[Schwartz et al., 1987] Schwartz, W. B., Patil, R. S., and Szolovits, P. (1987). Artificial intelligence in medicine, where do we stand? *Jurimetrics*, 27(4):362–369.

[Shaheen, 2021] Shaheen, M. Y. (2021). Applications of artificial intelligence (ai) in healthcare: A review. *ScienceOpen Preprints*.

[Shortliffe, 1974] Shortliffe, E. H. (1974). *MYCIN: A rule-based computer program for advising physicians regarding antimicrobial therapy selection.* PhD thesis, Stanford University Ph. D. dissertation.

[Tomsett et al., 2018] Tomsett, R., Braines, D., Harborne, D., Preece, A., and Chakraborty, S. (2018). Interpretable to whom? a role-based model for analyzing interpretable machine learning systems. *arXiv preprint arXiv:1806.07552*.

[Tschandl, 2021] Tschandl, P. (2021). Risk of bias and error from data sets used for dermatologic artificial intelligence. *JAMA dermatology*, 157(11):1271–1273.

[Uddin et al., 2023] Uddin, K. M. M., Biswas, N., Rikta, S. T., and Dey, S. K. (2023). Machine learning-based diagnosis of breast cancer utilizing feature optimization technique. *Computer Methods and Programs in Biomedicine Update*, 3:100098.

[Van der Velden et al., 2022] Van der Velden, B. H., Kuijf, H. J., Gilhuijs, K. G., and Viergever, M. A. (2022). Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis*, page 102470.

[Van Huffelen et al., 2024] Van Huffelen, A., Dijkgraaf, R., and Adriaansens, M. (2024). Aanbiedingsbrief (overheidsbrede) visie op generatieve ai.

[Van Woerkom, 2023] Van Woerkom, W. (2023). AFCBR. https://github.com/wijnanduu/AFCBR. Accessed: 2024-04-01.

[Van Woerkom et al., 2023] Van Woerkom, W., Grossi, D., Prakken, H., Verheij, B., Van Leeuwen, L., Verbrugge, R., and Renooij, S. (2023). Hierarchical a fortiori reasoning with dimensions. In Sileno, G., Spanakis, J., and van Dijck, G., editors, *Legal Knowledge and Information Systems. JURIX 2023: The Thirty-Sixth Annual Conference*, pages 43–52, Amsterdam etc. IOS Press.

[Veale and Binns, 2017] Veale, M. and Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, 4(2):2053951717743530.

[Wachter et al., 2020] Wachter, S., Mittelstadt, B., and Russell, C. (2020). Bias preservation in machine learning: the legality of fairness metrics under eu non-discrimination law. *West Virginia Law Review*, 123:735.

[Watson and Marir, 1994] Watson, I. and Marir, F. (1994). Case-based reasoning: A review. *The knowledge engineering review*, 9(4):327–354.

[Yeow et al., 2014] Yeow, W. L., Mahmud, R., and Raj, R. G. (2014). An application of case-based reasoning with machine learning for forensic autopsy. *Expert Systems with Applications*, 41(7):3497–3505.

[Zhang et al., 2022] Zhang, Y., Weng, Y., and Lund, J. (2022). Applications of explainable artificial intelligence in diagnosis and surgery. *Diagnostics*, 12(2):237.

9 Appendix

9.1 Preference relations (dimension: preferences)

ANGINA_PECTORIS__ANG_PECT_INSTAB: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
ANGSTSTOORNIS__ANGSTST_DATUM_DIAGN: Descending (-0.01)
BEWUSTZIJNSSTOORNIS_ECI__BEWUSTZIJNSST_ECI_LAATSTE_EP_2: [0]<3 JAAR (-0.01) $\preceq_1$ [1] NIET_AANW (0.01)
BIPOLAIRE_STOORNIS__BIP_STOORN_DATUM_DIAG: Ascending (0.0)
BLOEDGLUCOSE__BLOEDGLUC_NIET_NUCHTER: Descending (-0.1)
BLOEDGLUCOSE__BLOEDGLUC_NUCHTER_MMOL: Descending (-0.06)
CARDIALE_SYNCOPE__CARD_SYNC_PACEM: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
CATARACTEXTRACTIE__CATARACTEXTRACTIE_PSEUDO: [0] ODS (-1) $\preceq_1$ [1] OS (-1) $\preceq_1$ [2] OD (-1) $\preceq_1$ [3] NIET_AANW (1)
CNS__CNS_DIALYSE: [0] T (-0.03) $\preceq_1$ [1] F (0.03)
CNS__CNS_DM: [0] T (-0.03) $\preceq_1$ [1] F (0.03)
CNS__CNS_HARTRITMESTOORN: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
CNS__CNS_HART_EN_VAATZIEKTE: [0] T (-0.0) $\preceq_1$ [1] F (0.0)
CNS__CNS_IDOPATISCH: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
CNS__CNS_MDRD_2: [0] > 20 (-0.01) $\preceq_1$ [1] NIET_AANW (0.01)
CNS__CNS_MDRD_EGFR: Ascending (0.06)
CNS__CNS_MDRD_STAD_IV_15-19: [0] T (-0.02) $\preceq_1$ [1] F (0.02)
CNS__CNS_MDRD_STAD_IV_20_29: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
CNS__CNS_MDRD_STAD_I_II_III_30: [0] T (-0.0) $\preceq_1$ [1] F (0.0)
CNS__CNS_MDRD_STAD_V_KD_15: [0] T (-0.03) $\preceq_1$ [1] F (0.03)
CNS__CNS_NF_PERCENTAGE: Ascending (0.03)
COGNITIE__COGNITIE_MMSE: Ascending (0.01)
COGNITIE__COGNITIE_O: Descending (-0.01)
COGNITIE__COGNITIE_P: Descending (-0.01)
COGNITIE__COGNITIE_S: Descending (-0.01)
DEPRESSIE_BIPOLAIRE_STOORNIS__DEPRES_BIP_STOORN_LTST_EPIS_2: [0]< 5 JAAR (-0.01) $\preceq_1$ [1] > 5 JAAR (-0.0) $\preceq_1$ [2] NIET_AANW (0.01)
DEPRESSIE_BIPOLAIRE_STOORNIS__DEPRES_BIP_ST_BEH_PSYCH_AF_2: [0]<1 JAAR (-0.01) $\preceq_1$ [1] > 1 JAAR (0.01) $\preceq_1$ [2] NIET_AANW (0.01)
DEPRESSIE__DEPRES_BEH_PSYCH_1J: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
DEPRESSIE__DEPRES_DATUM_DIAG: Descending (-0.0)
DIPLOPIE__DIPLOPIE_NIET_HINDERLIJK: [0] F (nan)
DM__DM_ANTI_DIABETICA: [0] T (-0.03) $\preceq_1$ [1] F (0.03)
DM__DM_BEGINDATUM_GD_10_JR: [0] T (-0.03) $\preceq_1$ [1] F (0.03)
DM__DM_BEGINDATUM_GETAL: Ascending (0.38)
DM__DM_DATUM_LAATSTE_HYPO: Descending (-0.03)
DM__DM_DRP: [0] T (-0.03) $\preceq_1$ [1] F (0.03)
DM__DM_ERNSTIG_HYPO_1_JAAR: [0] T (-0.02) $\preceq_1$ [1] F (0.02)
DM__DM_GEEN_BEGRIP: [0] T (-0.02) $\preceq_1$ [1] F (0.02)
DM__DM_GEEN_CONTROLE: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
DM__DM_INSULINE: [0] T (-0.21) $\preceq_1$ [1] F (0.21)
DM__DM_NIET_INSULINE_HYPO_+: [0] T (-0.18) $\preceq_1$ [1] F (0.18)
DM__DM_NIET_INSULINE_HYPO_-: [0] T (-0.27) $\preceq_1$ [1] F (0.27)
DYSTHYMIE__DYSTHYMIE_DATUM_DIAGN: Ascending (0.01)
EPILEPTISCHE_AANVAL_LEN)__ANTI_EPILEP_STOPDATUM: Ascending (0.01)
EPILEPTISCHE_AANVAL_LEN)__EPILEP_AANVAL_LAATSTE_AANV_2: [0]< 5 JAAR (-0.01) $\preceq_1$ [1] > 5 JAAR (-0.0) $\preceq_1$ [2] NIET_AANW (0.0)

EPILEPTISCHE_AANVAL_LEN)__EPILEP_ANDERE_LAATSTE_AANV_DAT: Descending (-0.02)
EPILEPTISCHE_AANVAL_LEN)__EPILEP_EENM_AANVAL_DATUM: Descending (-0.01)
EPILEPTISCHE_AANVAL_LEN)__EPILEP_EENV_EERSTE_AANV_DAT: Descending (-0.01)
EPILEPTISCHE_AANVAL_LEN)__EPILEP_EENV_LAATSTE_AANV_DAT: Descending (-0.0)
EPILEPTISCHE_AANVAL_LEN)__EPILEP_MYO_EERSTE_AANV_DAT: Descending (-0.01)
EPILEPTISCHE_AANVAL_LEN)__EPILEP_MYO_LAATST_AANV_DAT: Descending (-0.0)
EPILEPTISCHE_AANVAL_LEN)__EPILEP_SLAAP_EERSTE_AANV_DAT: Descending (-0.01)
EPILEPTISCHE_AANVAL_LEN)__EPILEP_SLAAP_LAATSTE_AANV_DAT: Ascending (0.03)
EPILEPTISCHE_AANVAL_LEN)__EPILEP_SPOR_INT_GD_2_JAAR: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
EPILEPTISCHE_AANVAL_LEN)__EPI_AANVAL_WIJZ_AANGEP_MEDI: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
EPILEPTISCHE_AANVAL_LEN)__EPI_ANTI_EPILEPTICA: [0] T (-0.09) $\preceq_1$ [1] F (0.09)
EPILEPTISCHE_AANVAL_LEN)__EPI_EENMALIG_GEPROVOCEERD: Descending (-0.01)
FLAUWVALLEN__FLAUWVAL_DAT_LAATSTE: Descending (-0.01)
FLAUWVALLEN__FLAUWVAL_EENMALIG: Descending (-0.01)
FLAUWVALLEN__FLAUWVAL_MEERMALIG: [0] 2 (-0.01) $\preceq_1$ [1] NIET_AANW (0.01)
FUNCTIEBEP_BOVENSTE_LEDEMA__FUNCTIEBEP_BOV_HAND: [0] LINKS (-1) $\preceq_1$ [1] NIET_AANW (1)
FUNCTIEBEP_BOVENSTE_LEDEMA__FUNCTIEBEP_BOV_SCHOUD: [0] RECHTS (-1) $\preceq_1$ [1] LINKS (-1) $\preceq_1$ [2] NIET_AANW (1)
FUNCTIEBEP_BOVENSTE_LEDEMA__FUNCTIEBEP_BOV_VINGERS: [0] LINKS/RECHTS (-1) $\preceq_1$ [1] LINKS (0) $\preceq_1$ [2] RECHTS (0) $\preceq_1$ [3] NIET_AANW (1)
FUNCTIEBEP_ONDERSTE_LEDEMA__FUNCTIEBEP_OND_ENKEL: [0] RECHTS (-0.01) $\preceq_1$ [1] NIET_AANW (0.01)
FUNCTIEBEP_ONDERSTE_LEDEMA__FUNCTIEBEP_OND_KNIE: [0] RECHTS (-0.01) $\preceq_1$ [1] LINKS (-0.0) $\preceq_1$ [2] NIET_AANW (0.01)
FUNCTIEBEP_ONDERSTE_LEDEMA__FUNCTIEBEP_OND_VOET: [0] LINKS/RECHTS (-0.01) $\preceq_1$ [1] LINKS (-0.01) $\preceq_1$ [2] NIET_AANW (0.01)
FUNCTIEBEP_WERVELKOLOM__FUNCTIEBEP_WERV_CERVICAAL: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
FUNCTIEBEP_WERVELKOLOM__FUNCTIEBEP_WERV_LUMBAAL: [0] T (-0.0) $\preceq_1$ [1] F (0.0)
FUNCTIEBEP_WERVELKOLOM__FUNCTIEBEP_WERV_THORACAAL: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
GEBRUIK_RIJBEWIJS__GEBR_RIJB_CODE_101: [0] T (-0.04) $\preceq_1$ [1] F (0.04)
HARTKLEP_AFWIJKING__HARTKLEP_AFW_PROTHESE: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
HARTKLEP_AFWIJKING__HARTKLEP_INSUF_STEN: [0] T (-0.03) $\preceq_1$ [1] F (0.03)
HARTRITMESTOORNIS__HARTRITMESTOORN_BRADY: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
HARTRITMESTOORNIS__HARTRITMESTOORN_TACHY_AF/NSVT: [0] T (-0.0) $\preceq_1$ [1] F (0.0)
HART_EN_VAATAANDOENING__HART_VAAT_CCS: [0] II (-0.01) $\preceq_1$ [1] I (-0.0) $\preceq_1$ [2] NIET_AANW (0.0)
HART_EN_VAATAANDOENING__HART_VAAT_NYHA: [0] I (-0.03) $\preceq_1$ [1] II (-0.03) $\preceq_1$ [2] NIET_AANW (0.04)
HERSENDOORBLOEDINGSSTOORNIS__BEROERTE_DATUM: Ascending (0.01)
HERSENTUMOR__HERSENTUMOR_BEHANDELD: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
HERSENTUMOR__HERSENTUMOR_CUR_BEHAN: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
HERSENTUMOR__HERSENTUMOR_DATUM: Descending (-0.01)
HYPERTENSIE__HYPER_DIAS_METING: Descending (-0.04)
HYPERTENSIE__HYPER_SYS_METING: Descending (-0.09)
ICD__ICD_DATUM_LAATSTE_SHOCK: Ascending (0.04)
ICD__ICD_GEEN_UITDRAAI_AANW: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
ICD__ICD_IMPLANTATIEDATUM: Ascending (0.17)
ICD__ICD_PRIMAIR: [0] T (-0.13) $\preceq_1$ [1] F (0.13)
ICD__ICD_SECUNDAIR: [0] T (-0.11) $\preceq_1$ [1] F (0.11)
INTRAC_TUMOR_BUITEN_HERS__INTRAC_TUMOR_CUR_BEH: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
INTRAC_TUMOR_BUITEN_HERS__INTRAC_TUMOR_DATUM: Ascending (0.01)

MENIÈRE__MENIERE_LAATSTE_EPISODE_2: [0] $<$ 1 JAAR (-1) $\preceq_1$ [1] $>$ 1 JAAR (0) $\preceq_1$ [2] NIET_AANW (1)
MS__MS_BEGINDATUM: Descending (-0.01)
OOGAANDOENING__OOGAANDOEN_NYSTAGMUS: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
OOGAANDOENING__OOGAANDOEN_NYSTAGMUS_VERW: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
OOGAANDOENING__OOGAAND_ABL_RET_LIST: [0] ODS (-1) $\preceq_1$ [1] OOG NIET BEKEND (-1) $\preceq_1$ [2] OD (0) $\preceq_1$ [3] OS (0) $\preceq_1$ [4] NIET_AANW (1)
OOGAANDOENING__OOGAAND_OPT_PATHO: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
ORGAANTRANSPLANTATIE__HART: [0] T (-0.0) $\preceq_1$ [1] F (0.0)
ORGAANTRANSPLANTATIE__LEVER: [0] T (-0.0) $\preceq_1$ [1] F (0.0)
ORGAANTRANSPLANTATIE__LONGEN: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
ORGAANTRANSPLANTATIE__NIEREN: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
ORGAANTRANSPLANTATIE__ORGAANTRANS_DATUM: Ascending (0.01)
OSAS__OSAS_STARTDATUM_BEHAN: Ascending (0.05)
PROGRESSIEVE_OOGAANDOENING__PROGRESS_OOGAAN_MD: [0] ODS (-0.01) $\preceq_1$ [1] OS (-0.01) $\preceq_1$ [2] NIET_AANW (0.01)
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_CATARACT: [0] ODS (-1) $\preceq_1$ [1] OOG NIET BEKEND (-1) $\preceq_1$ [2] OD (0) $\preceq_1$ [3] OS (0) $\preceq_1$ [4] NIET_AANW (1)
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_DROGE_MD: [0] ODS (-0.09) $\preceq_1$ [1] OS (-0.05) $\preceq_1$ [2] OD (-0.04) $\preceq_1$ [3] NIET_AANW (0.11)
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_DRP: [0] ODS (-0.08) $\preceq_1$ [1] OD (-0.04) $\preceq_1$ [2] OS (-0.03) $\preceq_1$ [3] NIET_AANW (0.09)
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_GLAUCOOM: [0] ODS (-1) $\preceq_1$ [1] OOG NIET BEKEND (-1) $\preceq_1$ [2] OS (0) $\preceq_1$ [3] OD (0) $\preceq_1$ [4] NIET_AANW (1)
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_NATTE_MD: [0] OD (-0.04) $\preceq_1$ [1] ODS (-0.03) $\preceq_1$ [2] OS (-0.03) $\preceq_1$ [3] NIET_AANW (0.06)
PSYCHOSE_SCHIZOFRENIE__PSYCHOSE_SCHIZO_EENM_PSYCH_2: [0] $<$ 10 JAAR (-0.01) $\preceq_1$ [1] NIET_AANW (0.01)
PSYCHOSE_SCHIZOFRENIE__PSYCHOSE_SCHIZO_MEERDERE: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
REUMA__REUMA_ARTROSE: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
REUMA__REUMA_WEKE_DELEN: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
REUMA__REUMA__RA: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
SCHIZOFRENIESPECTRUMSTOORNIS__SCHIZOFRSPEC_STOORNIS: [0] T (-0.01) $\preceq_1$ [1] F (0.01)
VAD__VAD_IMPLANTATIEDATUM: Ascending (0.01)
VISUS__VISUS_VODS_MC: Ascending (0.13)
VISUS__VISUS_VODS_OC: Ascending (0.05)
VISUS__VISUS_VODS_ZC: Ascending (0.43)
VISUS__VISUS_VOD_MC: Ascending (0.15)
VISUS__VISUS_VOD_OC: Ascending (0.12)
VISUS__VISUS_VOD_ZC: Ascending (0.42)
VISUS__VISUS_VOS_MC: Ascending (0.11)
VISUS__VISUS_VOS_OC: Ascending (0.11)
VISUS__VISUS_VOS_ZC: Ascending (0.43)
ZIEKTE_VAN_PARKINSON__ZIEKTE_PARKINSON_BEGINDAT: Descending (-0.01)

9.2 Manual evaluation natures

The following natures were not taken into account in the configuration of the C-BR models, because of the requirement of manual evaluation of the driver fitness determination process:

- VRIJ_VELD_AANDOENING
- BTS
- HISTORIE
- MISBRUIK
- HERKEURING__HERKEU_BESL

9.3 Shorter validity period severities

The presence of the following severities suggest a shortened validity period of the license:

- GROEP1_KG1
- GROEP1_KG3
- GROEP1_KG5
- GROEP1_KG10
- GROEP1_ONGESCHIKT
- GROEP2_KG1
- GROEP2_KG3
- GROEP2_KG5
- GROEP2_ONGESCHIKT

9.4 Columns present in mistakes by AFR decisions

Column	Inc. forces	In testset	Ratio
REUMA__REUMA__RA	4	4	1,00
EPILEPTISCHE_AANVAL_LEN)__EPILEP_EENV_LAATSTE_AANV_DAT	1	1	1,00
ANGINA_PECTORIS__ANG_PECT_INSTAB	1	1	1,00
OOGAANDOENING__OOGAANDOEN_NYSTAGMUS_VERW	1	1	1,00
CNS__CNS_IDOPATISCH	1	1	1,00
ORGAANTRANSPLANTATIE__LEVER	1	1	1,00
HARTRITMESTOORNIS__HARTRITMESTOORN_BRADY	2	3	0,67
EPILEPTISCHE_AANVAL_LEN)__ANTI_EPILEP_STOPDATUM	10	19	0,53
CNS__CNS_NF_PERCENTAGE	7	14	0,50
BIPOLAIRE_STOORNIS__BIP_STOORN_DATUM_DIAG	1	2	0,50
REUMA__REUMA_WEKE_DELEN	1	2	0,50
OOGAANDOENING__OOGAAND_ABL_RET_LIST	3	7	0,43
HARTKLEP_AFWIJKING__HARTKLEP_AFW_PROTHESE	9	26	0,35
DM__DM_ERNSTIG_HYPO_1_JAAR	1	3	0,33
CNS__CNS_MDRD_STAD_IV_20_29	1	3	0,33
EPILEPTISCHE_AANVAL_LEN)__EPILEP_SPOR_INT_GD_2_JAAR	3	11	0,27
EPILEPTISCHE_AANVAL_LEN)__EPILEP_ANDERE_LAATSTE_AANV_DAT	8	30	0,27
COGNITIE__COGNITIE_O	1	4	0,25
COGNITIE__COGNITIE_P	1	4	0,25
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_GLAUCOOM	10	50	0,20
EPILEPTISCHE_AANVAL_LEN)__EPILEP_MYO_LAATST_AANV_DAT	1	5	0,20
EPILEPTISCHE_AANVAL_LEN)__EPILEP_EENM_AANVAL_DATUM	4	22	0,18
ORGAANTRANSPLANTATIE__ORGAANTRANS_DATUM	1	6	0,17
CNS__CNS_MDRD_EGFR	9	57	0,16
HART_EN_VAATAANDOENING__HART_VAAT_NYHA	3	20	0,15
HARTRITMESTOORNIS__HARTRITMESTOORN_TACHY_AF/NSVT	16	123	0,13
EPILEPTISCHE_AANVAL_LEN)__EPILEP_SLAAP_EERSTE_AANV_DAT	1	8	0,13
EPILEPTISCHE_AANVAL_LEN)__EPILEP_SLAAP_LAATSTE_AANV_DAT	1	8	0,13
HERSENDOORBLOEDINGSSTOORNIS__BEROERTE_DATUM	13	106	0,12
CATARACTEXTRACTIE__CATARACTEXTRACTIE_PSEUDO	56	480	0,12
DM__DM_ANTI_DIABETICA	1	9	0,11
DM__DM_BEGINDATUM_GD_10_JR	1	9	0,11
EPILEPTISCHE_AANVAL_LEN)__EPI_ANTI_EPILEPTICA	4	44	0,09
HARTKLEP_AFWIJKING__HARTKLEP_INSUF_STEN	1	11	0,09
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_CATARACT	27	304	0,09
HYPERTENSIE__HYPER_SYS_METING	4	46	0,09
VISUS__VISUS_VODS_OC	25	311	0,08
VISUS__VISUS_VOS_OC	24	318	0,08
VISUS__VISUS_VOD_OC	24	321	0,07
HYPERTENSIE__HYPER_DIAS_METING	1	14	0,07

Column	Inc. forces	In testset	Ratio
BLOEDGLUCOSE__BLOEDGLUC_NIET_NUCHTER	6	87	0,07
VISUS__VISUS_VODS_ZC	145	2222	0,07
VISUS__VISUS_VODS_MC	109	1732	0,06
VISUS__VISUS_VOD_MC	132	2099	0,06
VISUS__VISUS_VOS_MC	132	2101	0,06
VISUS__VISUS_VOS_ZC	166	2700	0,06
VISUS__VISUS_VOD_ZC	166	2701	0,06
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_DRP	2	44	0,05
DM__DM_INSULINE	8	197	0,04
DM__DM_BEGINDATUM_GETAL	22	601	0,04
BLOEDGLUCOSE__BLOEDGLUC_NUCHTER_MMOL	1	30	0,03
DM__DM_NIET_INSULINE_HYPO_+	5	156	0,03
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_NATTE_MD	1	34	0,03
DM__DM_NIET_INSULINE_HYPO_-	9	333	0,03
PROGRESSIEVE_OOGAANDOENING__PROGRES_OOGAAN_DROGE_MD	2	78	0,03

9.5 Health declaration

Uw gezondheid *(vervolg)*

Gedragsaandoening en psychiatrische diagnose

→ *Wilt u de hokjes die u kiest ■ helemaal opvullen?*

14 **Heeft u ADHD of ADD?**	☐ Ja	☐ Nee
15 **Heeft een arts bij u autisme of een vorm daarvan vastgesteld, zoals PDD-NOS, McDD of Asperger?**	☐ Ja	☐ Nee
16 **Heeft een arts een andere psychische of psychiatrische diagnose bij u gesteld?**	☐ Ja → *Vul ook vraag 16a t/m 16d in.*	☐ Nee → *Ga verder met vraag 17.*
16a Heeft een arts bij u een depressie of een bipolaire stoornis vastgesteld?	☐ Ja → *Vul ook de volgende vraag van 16a in.*	☐ Nee → *Ga verder met vraag 16b.*
Is de diagnose in de afgelopen vijf jaar gesteld?	☐ Ja → *Vul ook de volgende vraag van 16a in.*	☐ Nee → *Ga verder met vraag 16b.*
Bent u in het afgelopen jaar behandeld door een psychiater?	☐ Ja	☐ Nee
16b Heeft u wel eens een psychose gehad of heeft u schizofrenie?	☐ Ja → *Vul ook de volgende vraag van 16b in.*	☐ Nee → *Ga verder met vraag 16c.*
Heeft u tot nu toe maar één keer een psychose gehad?	☐ Ja → *Vul ook de volgende vraag van 16b in.*	☐ Nee → *Ga verder met vraag 16c.*
Heeft u die éénmalige psychose in de afgelopen 10 jaar gehad?	☐ Ja	☐ Nee
16c Heeft u een angststoornis?	☐ Ja	☐ Nee
16d Heeft u een andere psychische of psychiatrische aandoening die in de vorige vragen niet genoemd is?	☐ Ja	☐ Nee

Misbruik van middelen

17 **Heeft u de afgelopen vijf jaar overmatig gebruik (misbruik) gemaakt van alcohol, drugs of andere verslavende middelen?**	☐ Ja	☐ Nee

Medicijnen

18 **Gebruikt u medicijnen die de rijvaardigheid negatief beïnvloeden?**	☐ Ja	☐ Nee

Verdere klachten of aandoeningen

19 **Heeft u verder nog klachten of aandoeningen die u kunnen beperken bij het besturen van een voertuig?**	☐ Ja	☐ Nee

Ondertekening

Ik heb deze eigen verklaring over mijn gezondheid naar waarheid ingevuld.

Handtekening

Datum ☐☐ - ☐☐ - ☐☐☐☐
→ *Vul in als dd-mm-jjjj.*

Gezondheidsverklaring
Rijbewijs

cbr

Waarom vult u dit formulier in?
U wilt een rijbewijs aanvragen of verlengen. Of u wilt graag dat het CBR onderzoekt of u nog veilig kunt rijden met het rijbewijs dat u nu heeft.
Om te mogen rijden moet u 'rijgeschikt' zijn. Dat wil zeggen: geestelijk en lichamelijk gezond genoeg om te rijden. Dit formulier invullen is de eerste stap in de beoordeling van uw rijgeschiktheid. Het is belangrijk dat u de vragen eerlijk invult. Voor uw eigen veiligheid en de veiligheid van andere weggebruikers.

Wat moet u doen?
- Vul uw persoonlijke gegevens in. En kruis aan welk rijbewijs u wilt.
- Beantwoord de vragen over 'Uw gezondheid' en zet uw handtekening eronder.
- Stuur het formulier in de antwoordenvelop op naar het CBR.
- Binnen vier weken krijgt u van ons bericht.

Uw gegevens
U geeft persoonsgegevens aan ons door. Wij gebruiken deze voor uw aanvraag en daarmee samenhangende doelen.

Als wij u doorverwijzen naar een medisch specialist, dan krijgt die specialist de noodzakelijke medische gegevens van ons. Daarnaast geven wij rijbewijsgegevens door aan de gemeente en de Rijksdienst voor het Wegverkeer (RDW). Meer informatie vindt u in onze privacyverklaring op cbr.nl.

Heeft u vragen?
Op cbr.nl vindt u meer informatie.
U kunt ook onze klantenservice bellen: 088 227 77 00.

Uw persoonlijke gegevens

→ *Wilt u de vragen met een zwarte of blauwe pen invullen?*

Voorletters en achternaam	*Voorletters* / *Achternaam* → *Vul uw volledige achternaam in.*
Adres → *Vul uw postadres in.*	*Straatnaam* / *Huisnummer* / *Toevoeging* *Postcode* / *Plaatsnaam*
Burgerservicenummer (BSN)	→ *U vindt dit nummer op uw paspoort, identiteitskaart of zorgpas.*
Geboortedatum	→ *Vul in als dd-mm-jjjj.* ☐☐ - ☐☐ - ☐☐☐☐
Telefoonnummer overdag	
E-mailadres	

Uw rijbewijs

→ *Wilt u de hokjes die u kiest ■ helemaal opvullen?*

Welke rijbewijzen wilt u op uw nieuwe rijbewijs hebben? Kruis de rijbewijzen aan die u nu heeft en wilt houden. En het rijbewijs waar u examen voor gaat doen.

Voor het rijbewijs Bromfiets, snorfiets of brommobiel (AM) is geen gezondheidsverklaring nodig.

Soort rijbewijs

- ☐ Motor of motorscooter **(A1, A2, A)**
- ☐ Personenauto en bestelauto **(B)**
- ☐ Aanhangwagen achter auto **(BE)**
- ☐ Tractor of motorvoertuig met beperkte snelheid **(T)**

→ *Ga naar de volgende pagina.*

Uw gezondheid

→ *Wilt u de hokjes die u kiest ■ helemaal opvullen?*

Armen en benen

1	**Kunt u uw arm, hand of vingers door een beperking slecht gebruiken?**	☐ Ja	☐ Nee
2	**Kunt u uw been of voet door een beperking slecht gebruiken?**	☐ Ja	☐ Nee

Zicht en ogen

3	**Kunt u met één of beide ogen beperkt zien, zelfs als u een bril of lenzen gebruikt?**	☐ Ja → *Vul ook vraag 3a in.*	☐ Nee → *Ga verder met vraag 4.*
3a	Rijdt u of wilt u gaan rijden met een bioptische telescoop (BTS) via VISIO?	☐ Ja	☐ Nee
4	**Wordt of werd u behandeld door een oogarts voor iets anders dan voor een bril of lenzen?**	☐ Ja → *Vul ook vraag 4a in.*	☐ Nee → *Ga verder met vraag 5.*
4a	Staat u nu nog steeds onder controle van een oogarts?	☐ Ja	☐ Nee

Diabetes mellitus

5	**Heeft u diabetes mellitus?**	☐ Ja → *Vul ook vraag 5a in.*	☐ Nee → *Ga verder met vraag 6.*
5a	Gebruikt u op dit moment medicijnen voor de diabetes?	☐ Ja → *Vul ook de volgende vraag van 5a in.*	☐ Nee → *Ga verder met vraag 6.*
	Heeft u de diabetes al langer dan 10 jaar?	☐ Ja	☐ Nee

Verminderde nierfunctie

6	**Heeft u chronische schade aan uw nieren (een verminderde nierfunctie)?**	☐ Ja → *Vul ook vraag 6a in.*	☐ Nee → *Ga verder met vraag 7.*
6a	Is uw nierfunctie minder dan 20%?	☐ Ja	☐ Nee

Aandoening van longen of bloed, transplantatie

7	**Heeft u een aandoening van uw longen (COPD), een hoge bloeddruk of een bloedziekte? Of heeft u een orgaantransplantatie gehad?**	☐ Ja	☐ Nee

Hart en bloedvaten

8	**Heeft u een ziekte van uw hart of van uw bloedvaten?**	☐ Ja → *Vul ook vraag 8a en 8b in.*	☐ Nee → *Ga verder met vraag 9.*
8a	Heeft u een ICD?	☐ Ja	☐ Nee
8b	Heeft u een steunhart?	☐ Ja	☐ Nee

→ *Ga naar de volgende pagina.*

Aandoening van het zenuwstelsel

9	**Heeft een arts bij u een aandoening vastgesteld van uw ruggenmerg, zenuwstelsel of hersenen (bijvoorbeeld een beroerte of spierziekte)?**	☐ Ja → *Vul ook vraag 9a t/m 9g in.*	☐ Nee → *Ga verder met vraag 10.*
9a	Heeft een arts bij u een vorm van dementie of MCI vastgesteld?	☐ Ja	☐ Nee
9b	Heeft u een beroerte, een herseninfarct of een hersenbloeding gehad?	☐ Ja → *Vul ook de overige vragen van 9b in.*	☐ Nee → *Ga verder met vraag 9c.*
	Is dit in de afgelopen zes maanden gebeurd?	☐ Ja	☐ Nee
	Bent u aan de bloedvaten van uw hersenen geopereerd?	☐ Ja	☐ Nee
	Heeft u nog steeds klachten van uw beroerte of herseninfarct?	☐ Ja	☐ Nee
9c	Is er bij u een misvorming van een hersenbloedvat ontdekt, zonder dat u eerst een bloeding had?	☐ Ja	☐ Nee
9d	Heeft u een ziekte van uw zenuwstelsel?	☐ Ja	☐ Nee
9e	Heeft u een spierziekte?	☐ Ja	☐ Nee
9f	Heeft u een hersentumor? Of heeft u er een gehad?	☐ Ja	☐ Nee
9g	Heeft u een aandoening van uw hersenen of zenuwstelsel die nog niet genoemd is?	☐ Ja	☐ Nee

Epilepsie, slaperigheid en bewusteloosheid

10	**Heeft u wel eens een epileptische aanval gehad?**	☐ Ja → *Vul ook vraag 10a in.*	☐ Nee → *Ga verder met vraag 11.*
10a	Heeft u de laatste epileptische aanval in de afgelopen vijf jaar gehad?	☐ Ja	☐ Nee
11	**Heeft u wel eens overdag last gehad van abnormale slaperigheid?**	☐ Ja → *Vul ook vraag 11a en 11b in.*	☐ Nee → *Ga verder met vraag 12.*
11a	Heeft een arts slaapapneu (OSAS) bij u vastgesteld?	☐ Ja	☐ Nee
11b	Heeft een arts narcolepsie of hypersomnolentie bij u vastgesteld?	☐ Ja	☐ Nee
12	**Bent u in de afgelopen drie jaar meerdere keren plotseling uw bewustzijn verloren (bijvoorbeeld flauwgevallen)?**	☐ Ja	☐ Nee

Ziekte van Menière

13	**Heeft u als gevolg van de ziekte van Menière in het afgelopen jaar aanvallen van duizeligheid gehad?**	☐ Ja	☐ Nee

→ *Ga naar de volgende pagina.*