Towards a Case-Based Decision Support System for Recruiting Processes using T-Shapes

Jakob Michael Schoenborn[1,3], Pascal Reuss[1,3], Carsten Wenzel[2], Klaus-Dieter Althoff[1,3]

Abstract: Modeling the individual set of skills a person possesses is a difficult task. However, most people seek to find a job that suits them best, despite being unaware of which skills she inherits and what defines herself as a person. Recruiters are knowledgeable in terms of extracting these information by an informal phone call or when carrying out an assessment center. However, we suggest to set a step beforehand and provide a case-based decision support system by using T-Shapes as a model for matching personal skills with job requirements. We provide a first overview on how to structure and model the four knowledge containers of case-based reasoning and how we use these knowledge containers to obtain the most similar T-Shape (situation). This T-Shape can then be reused by the recruiter, providing her with all necessary information at a first glance.

Keywords: Case-based Reasoning, Decision Support, T-Shapes, Recruiting

1 Introduction

The process of finding the most suitable occupation can often be daunting, given the huge range of different possible career paths. Especially for students who are about to finish their compulsory education, it is often unclear to them which occupation suits them the best, given their individual interests and skills. Schools and universities are working together with career services to address this challenge, for example by carrying out assessment centers. However, only 17 % of 2010 to 2016 graduates of the Purdue University (USA) report that these were helpful as the Gallup-Purdue index study shows [Ga16]. On a similar note, only 15 % of German employees report an emotional commitment to their working places [Ni18], resulting in 85 % employees to be likely to change their current occupation given the right opportunity. These different audiences are the targets for the recruiting domain in general. For the last years, and especially in the STEMM (Science, Technology, Engineering, Mathematics, and Medicine) domains, the popularity of headhunting increased. The majority of large German companies rely on headhunting to find the right applicants for the right position [FH17, Mi00]. Headhunting itself is not a new method in terms of recruiting. However, due to the rising popularity and especially possible range of platforms like XING[4] and LinkedIn[5] (among others), individuals are able to present themselves and

[1] Intelligent Information Systems Lab, University of Hildesheim, Hildesheim, Germany
[2] Department of Information Systems and Enterprise Modelling, University of Hildesheim, Hildesheim, Germany
[3] Competence Center Case Based Reasoning, German Center for Artificial Intelligence, Kaiserslautern, Germany
[4] URL: https://www.xing.com
[5] URL: https://www.linkedin.com/

their skills to the publicity, providing the headhunters with the required initial information [FH17].

In the recent literature, there is an uprising interest in automating this process or at least support the so-called screening phase of the recruitment process and to use the digitization to decrease the requirement of paper application management. Espenakk et al. presented a lazy learning approach by considering user profiles as feature vectors using attributes such as occupation, skill, language among others combined with case-based reasoning (CBR) [EKK19]. Earlier stages of recruitment using CBR has been contributed by Siraj et al. [Si11] and Rafter et al. [RBS00]. Another feature vector approach by using machine learning techniques can be seen by Faliagka et al., solving the candidate ranking problem using LinkedIn profiles and linguistic analysis [Fa12].

The mapping between the skills of a possible applicant to the current open job position is still challenging. To tackle this challenge, we suggest the usage of T-Shapes. These T-Shapes are considering two different dimensions: vertical for skills with rather deepened knowledge; and horizontal for broader skills, for example, where the applicant has gathered experiences but would not be considered as an expert (unlike vertical skills). An example of T-Shapes can be seen in Fig. 2. T-Shapes have been introduced by David Guest in 1991 [Gu91], but received more attention after Tim Brown, CEO of IDEO (a global design company) mentioned using these models. The contents (attributes) of the T-Shape can either be manually inserted by the person herself (the recruiter) or can be gathered in form of the result of an assessment center. Some larger companies require to pass an assessment center before the application process itself begins. The results of these assessment centers aim to show strengths and weaknesses of considered applicants, which in turn could be transferred into T-Shapes. We propose to use these models as case structures for the methodology of CBR.

Case-based reasoning uses experiences of past situations, stores them in a case base, and reuses the stored experiences for future, similar problems. The so-called CBR-cycle (see Fig. 1) consists of four steps: Retrieve, Reuse, Revise, Retain and is accompanied by four knowledge containers: similarity measurements, vocabulary, case base and adaptation knowledge [Ba12, Ri03].

The first step, retrieve, uses the input situation of a user - the current situation. Based on the attributes and their values which describe the situation, the best matching stored case from the case base will be retrieved, using similarity measurements. During the second step, the retrieved case will be proposed to be reused, thus, provided to the user. The user might accept this solution or reject it due to various reasons, e.g., the similarity between the input and the retrieved situation is not high enough. Either way will trigger the next step, revise, where the proposed case or the knowledge containers can be adjusted to provide a tested or repaired case to the user. During the last step, retain, the knowledge engineer may decide to store the case for further usage, or to store adjustments of the similarity measurements to

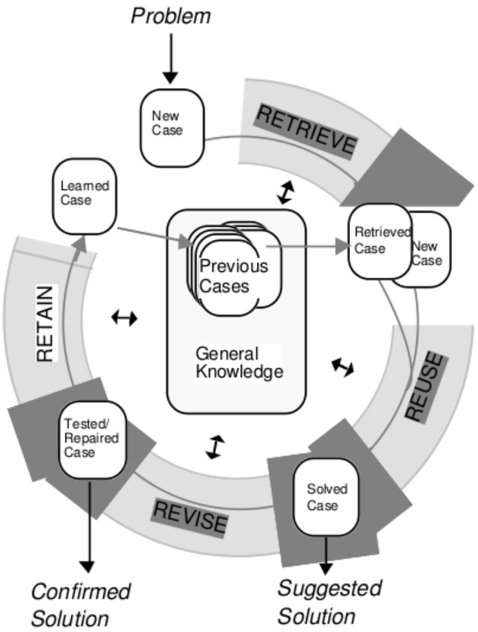

Fig. 1: Case-based reasoning (CBR)-cycle by A. Aamodt and E. Plaza [AP94].

prevent the case base from growing too large, running into a risk of redundancy or lowered maintainability.

2 Using T-Shapes for domain modeling in CBR

As mentioned above, we suggest the combination of T-Shapes and CBR. Fig. 2 depicts an exemplary T-Shape of a person considering herself as very knowledgeable in case-based reasoning. Nevertheless, other areas such as penetration testing and cryptography are also a feature of the skill set of this person, for example by possessing certain certificates. However, this model would suggest that the person considers her own main interest and main contribution in the area of case-based reasoning. Regarding the attribute-value pairs and the suggested granularity as well as the scope of the used attributes, we consider T-Shapes as a simplistic depiction of a certain real world application which has been developed to serve a concrete purpose. It reduces the complexity of the original domain to generalise information in a way which is simple to understand and, thus, enables different users to communicate about complex problems. T-Shapes are an example of a model due to its simplified depiction of a persons set of skills. Using this intentionally broad definition of a T-Shape, the emphasis is on simplifying the very complex domain of competence management. Dealing with every facet of competence management is a complex task whereas the possibility to apply a first

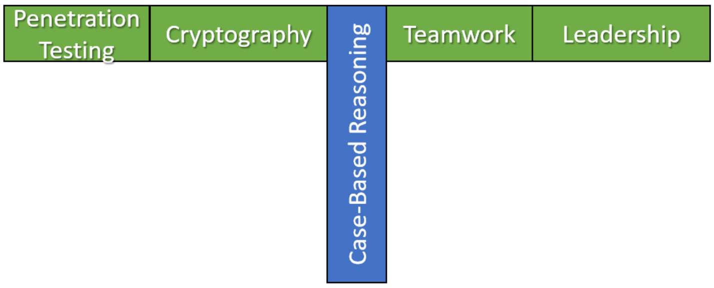

Fig. 2: Examplary T-Shape model showing the skill set of a person.

filter and then moving further into detail is expected to be beneficial regarding time until a final decision (not) to recruit has been made. We use this simplification to store T-Shapes as cases in the case base to ensure an efficient retrieval process in terms of computational time while still considering complex problems by using the knowledge containers.

For each attribute in the horizontal domain, the width of each skill represents the knowledge level of the corresponding skill: the larger the width, the more deepened the skill. This opens up the question on the model itself: Would not only one dimension be enough whereas the largest container represents in which skill the user can be considered as an expert? Indeed, there are multiple different variations of shapes, such as the I-Shape (which would represent the mentioned idea), the π-Shape, as well as the M-Shape (which suggests two/three depth-skills, which can be seen analogous to the top 3 skills a premium member can name on XING). However, the depth skill does not necessarily needs to be limited to one skill as suggested in Fig. 2 but rather can split up into multiple sub-domains, for example, CBR regarding maintenance, knowledge management or gaming. Nevertheless, it is important to list it as *one* attribute to emphasize the main interest area of an applicant and to categorize appropriately during the first fielding step.

Skills can be divided into different categories, for example: Technical skills and social skills. The former are usually easier to quantify in terms of certificates, number of completed projects, and years of experience. However, as also often listed in job positions, social skills such as being able to work in a team but also to be self-reliant are also required. Nevertheless, experience made by hobbies or side-jobs are important criteria of a persons personality and thus need to be contained in the T-Shape model. An exemplary approach can be seen by the work of Wilkin et al. who mapped activities like “has supervised the work of others” to scores for the five attributes “Intelligence, Interpersonal, Leadership, Motivation, Usefulness” whereas Leadership scored as the highest [WC12]. By comparing the complexity of soft skills in contrast to hard skills, there is a higher complexity caused by dependencies between multiple soft skills recognizable. As an example we consider the skill of problem solving thinking: on the one hand, this can be regarded as an independent skill but also requires a lot of other skills, such as the ability to grasp complex interrelationships, the ability to plan, and the ability to think and act in a complex way. By using completion rules, however, we are able to complete requests in cases where more complex soft skills require lower soft skills.

The IT skill pyramid, for example, is used to classify the dependency on skills [PT17]. It divides skills into three tiers: The lowest tier describes the fundamental knowledge needed to start a career in the IT field, including for example the above mentioned problem solving thinking. Additionally, general topics like the knowledge about the usage of certain established technologies are covered in the first tier. Tier two contains all the technical skills, knowledge and abilities required to work in the IT field, such as an understanding of the software and hardware and the underlying architecture. The third level describes the skills that require a comprehensive overview and knowledge of the IT sector [PT17]. Hard skills are mostly unique, such as graduation, university degrees and certificates.

Within T-Shapes, the knowledge is available in an unstructured form. In order to use this knowledge in CBR, we have to transfer it into the structured form of CBR in hindsight of the four knowledge containers as introduced. We propose the following structure of a case C containing a vertical, most weighted, skill V, a set of horizontal skills H containing technical- and soft-skills, and a job position P as solution to the case:

$$C = (V, \{H\}, P) \tag{1}$$

As soon as each attribute has been structured or mapped into a measurable way by using *similarity measurements*, for example, by numeric functions, symbolic matrix representations, Boolean comparisons or taxonomies of social skills (as described above regarding the matching of soft-skills), we can fill the case base with cases of individual persons and another case base with cases of job advertisements to compare against. Within the *vocabulary*, the structure of the used model is defined as well as all values that may belong to these attributes. Additionally, it is important to store synonyms to provide a proper matching between skills and solutions of the case, for example, by using a thesaurus with synonymous of job positions[6].

The knowledge container of *adaptation knowledge* contains information on how a T-Shape can be used to solve a query case. As a concrete example, a person with a master degree in meteorology has a high similarity to a data analyst due to her personality profile. The skills of recognizing patterns and dealing with large amounts of data are among the top qualifications for both persons and positions. From this point, it could be deduced that a meteorologist might also be suitable for the position of a data analyst. The reverse conclusion, however, is not always granted. Due to the lack of domain knowledge about meteorology, a trained data analyst will not be able to make reliable predictions about the weather forecast.

In the *case base*, known T-Shapes with their corresponding job descriptions are stored. All existing attributes from the job description as well as created T-Shapes are converted into attribute-value pairs. This differentiation and way of modelling allows us the comparison

[6] e.g., as provided by John Carty: https://github.com/johnpcarty/Thesaurus-of-Job-Titles

between two T-Shapes, thus, enabling to highlight the most similarity between those two. Additionally, we can also provide the most dissimilar case. By enabling the user to provide feedback, we may be able to adjust the similarity measurements (as one of the knowledge containers) to increase the accuracy in general and, in the best case, create completion rules to also enable the system to provide a good solution, despite a barely filled T-Shape.

3 Conclusion and Outlook

We provided a first glance on our approach to support recruiters and individual persons by finding an accurate job position. Due to the early stage of progress, it should be clearly noted that an improvement of our domain model is currently in progress. A far-reaching vocabulary needs to be created automatically with the help of text mining methods. After the refinement of the vocabulary, the newly acquired attribute values have to be entered into our similarity measures. This can be done with a framework like FEATURE-TAK (**F**ramework for **E**xtraction, **A**nalysis, and **T**ransformation of **U**nstructu**RE**d **T**extual **A**ircraft **K**nowledge) that extends knowledge containers of a CBR system in a semi-automated way. The framework was originally developed to analyze, extract, and transform unstructured or semi-structured textual knowledge from failure descriptions in the aviation domain. The extracted knowledge items then were used to extend the vocabulary, the similarity measures, and the case bases of a decision support system for diagnosis and maintenance of Airbus aircraft [Re18]. While it was developed for textual aircraft knowledge, the framework can be extended and configured to be used for other textual knowledge, too. The creation of a further prototype and its evaluation is also planned to gain feedback of the application and to measure the users acceptance in terms of allowing algorithmic decision making taking place in crucial decisions of an individual person regarding her individual development and career path.

Bibliography

[AP94] Aamodt, Agnar; Plaza, Enric: Case-based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. AI Commun., 7(1):39–59, March 1994.

[Ba12] Bach, Kerstin: Knowledge Acquisition for Case-Based Reasoning Systems. PhD thesis, University of Hildesheim, 2012.

[EKK19] Espenakk, Erik; Knalstad, Magnus Johan; Kofod-Petersen, Anders: Lazy Learned Screening for Efficient Recruitment. In: Case-Based Reasoning Research and Development - 27th International Conference, ICCBR 2019, Otzenhausen, Germany, September 8-12, 2019, Proceedings. pp. 64–78, 2019.

[Fa12] Faliagka, Evanthia; Ramantas, Kostas; Tsakalidis, Athanasios; Tzimas, Giannis: Application of Machine Learning Algorithms to an online Recruitment System. 01 2012.

[FH17] Flöthmann, Christoph; Hoberg, Kai: Career Patterns of Supply Chain Executives: An Optimal Matching Analysis. Journal of Business Logistics, 38, 2017.

[Ga16] Gallup: Great Jobs. Great Lives. The Value of Career Services, Inclusive Experiences and Mentorship for College Graduates. 2016. Last validation: 11/28/2019.

[Gu91] Guest, David E.: Personnel Management: The End of Orthodoxy? British Journal of Industrial Relations, 29(2):149–175, 1991.

[Mi00] Mileham, Patrick: The 'science' of headhunting. Drug Discovery Today, 5(4):161 – 163, 2000.

[Ni18] Nink, Marco: Engagement Index Germany 2018. 2018.

[PT17] Patacsil, Frederick; Tablatin, Christine: Exploring the importance of soft and hard skills as perceived by IT internship students and industry: A gap analysis. Journal of Technology and Science Education, 7:347, 09 2017.

[RBS00] Rafter, Rachael; Bradley, Keith; Smyth, Barry: Personalised Retrieval for Online Recruitment Services. 04 2000.

[Re18] Reuss, Pascal; Stram, Rotem; Althoff, Klaus-Dieter; Henkel, Wolfram; Henning, Frieder: Knowledge Engineering for Decision Support on Diagnosis andMaintenance in the Aircraft Domain. In (Nalepa, Grzegorz J.; Baumeister, Joachim, eds): Synergies Between Knowledge Engineering and Software Engineering, volume 626. Springer, pp. 173–196, advances in intelligent systems and computing edition, 2018.

[Ri03] Richter, Michael M.: Fallbasiertes Schließen. Görz, Günther; Rollinger, Claus-Rainer; Schneeberger, Josef (Hrsg.): Handbuch der Künstlichen Intelligenz, 4:407–430, 2003.

[Si11] Siraj, Fadzilah; Mustafa, Norhashimah; Haris, Megat; Yusof, Shahrin; Muhammad Ashraq, Salahuddin; Hasan, Md Rajib: Pre-selection of Recruitment Candidates Using Case Based Reasoning. Computational Intelligence, Modelling and Simulation, International Conference on, 0:84–90, 09 2011.

[WC12] Wilkin, Christa L.; Connelly, Catherine E.: Do I Look Like Someone Who Cares? Recruiters' Ratings of Applicants' Paid and Volunteer Experience. International Journal of Selection and Assessment, 20(3):308–318, 2012.