Personalised Learning Environments based on Knowledge Graphs and the Zone of Proximal Development

Yoshi Malaise\(^a\) and Beat Signer\(^b\)

Web & Information Systems Engineering Lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

Keywords: Personalised Learning, Knowledge Graph, Personalised Teaching Assistant Tool, e-Learning, Smart Learning.

Abstract: The learning of new knowledge and skills often requires previous knowledge, which can lead to some frustration if a teacher does not know a learner’s exact knowledge and skills and therefore confronts them with exercises that are too difficult to solve. We present a solution to address this issue when teaching techniques and skills in the domain of table tennis, based on the concrete needs of trainers that we have investigated in a survey. We present a conceptual model for the representation of knowledge graphs as well as the level at which individual players already master parts of this knowledge graph. Our fine-grained model enables the automatic suggestion of optimal exercises in a player’s so-called zone of proximal development, and our domain-specific application allows table tennis trainers to schedule their training sessions and exercises based on this information. In an initial evaluation of the presented solution for personalised learning environments, we received positive and promising feedback from trainers. We are currently investigating how our approach and conceptual model can be generalised to some more traditional educational settings and how the personalised learning environment might be further improved based on the expressive concepts of the presented model.

1 INTRODUCTION

Learning the ropes of a new activity can often be a daunting task, in particular if the learning material requires a lot of prior knowledge. Situations like these often lead to frustration since the learner has no way to figure out the exact skills and knowledge they are missing in order to properly advance on their path. Likewise, for educators and coaches it can be difficult to provide proper guidance if they do not know their students’ past learning trajectory and their current proficiency levels.

In previous research, the use of knowledge graphs has been proposed as a way to provide a semantic representation of all relevant topics and their relations for a given domain (Rizun, 2019). In such a knowledge graph the topics are represented as nodes of a directed acyclic graph with the edges representing specific associations. These associations typically indicate that a certain topic should be explored first, as the knowledge and skills learned in that topic are necessary to master a more advanced topic. The formal representation of topics can be used to provide an overview of which topics a learner could study next. Furthermore, by automatically navigating the graph during diagnostic assessments, we can detect a learner’s knowledge gaps. In more structured environments with clear requirements to reach at intermediate milestones (e.g. the six year curriculum to obtain a high school diploma is split into a different set of learning objectives for each two-year interval), a predefined walk through the knowledge graph in the form of a so-called learning path can be defined. When students move from one curriculum to another—for instance after a relocation—the knowledge paths of the two curricula can be compared to provide assistance during the transition period (Ilkou and Signer, 2020).

We believe that there are great opportunities for adaptive and personalised learning environments taking a user’s zone of proximal development into account. Our goal is to use knowledge graphs in combination with a user’s acquired skills to automatically detect and recommend the topics to be learned next in the zone of proximal development. We present a tool to assist Flemish table tennis trainers in teaching the fundamental skills of their sport. It serves as a concrete personalised learning environment as well as a proof of concept to explore how the approach could be generalised to a broader learning context.
We start by highlighting some related work and introducing the concept of proximal development. Based on a literature study and a survey that we conducted to find out more about the needs of trainers teaching fundamental table tennis skills, we derive a number of requirements for a personal learning environment solution. After presenting a prototype of such a personal learning environment based on knowledge graphs and the zone of proximal development, we discuss its generalisation to other domains and provide some general conclusions.

2 BACKGROUND

Knowledge graphs have, for instance, been used to add some advanced navigation to the Moodle\(^1\) learning management system (Scherl et al., 2012), resulting in an increase in participants’ overall knowledge compared to Moodle’s classical interface.

A key construct of Vygotski’s theory of learning is the identification of the zone of proximal development (Vygotsky, 1978). The zone is defined as the space between what a learner can do unaided and the tasks that are too difficult to be performed even under some guidance, as illustrated in Figure 1. Tasks in this zone are optimal to improve a learner’s knowledge, as they are challenging enough to push the learner, but not impossible to master.

![Figure 1: Zone of proximal development.](image)

We believe that enriching educational content such as explanations, exercises and multimedia content with knowledge graphs can form the core of a smart, adaptive and personalised learning environment, always suggesting exercises in a user’s zone of proximal development. Thereby, less time is “wasted” in performing assignments covering already known topics and we can reduce the risk of demotivating learners and lowering their confidence by presenting them too difficult exercises (Shabani et al., 2010).

Similar research has recently been conducted based on data made available via the Learnta TAD\(^2\) online platform (Zou et al., 2019; Baker et al., 2020). As a result of their evaluation of k-12 students in the fields of Mathematics and English, the authors concluded that in both fields the students managed to complete more tasks if the tasks were selected from their zone of proximal development. A major difference to our proposed solution is that they used statistical methods to estimate the mastery of topics whereas we aim for a teacher-in-the-loop approach.

In order to properly define the requirements for a personalised learning tool for the field of table tennis, we relied on information provided to us by trainers in a survey, as well as some literature research on the use of knowledge graphs in education and models used in sports education such as the three-stage model of motor skill acquisition (Fitts and Posner, 1967) and Kolb’s experiential learning cycle (Kolb, 2014).

![Figure 2: Three-stage model of motor skill acquisition.](image)

Kolb’s experiential learning cycle theory—with its four-stage cycle of learning illustrated in Figure 3—emphasises the importance of a learner’s experiences while acquiring a new skill (Kolb, 2014). In the active experimentation stage, the learner tries to perform an action and experiences the result of their action (success or failure) in the concrete experience.

![Figure 3: Kolb’s experiential learning cycle.](image)

\(^{1}\)https://moodle.org  
\(^{2}\)https://www.learnta.com
stage. A learner then enters the reflective observation stage where they reflect about their action and its result, potentially assisted by feedback from a supervisor. The learner further moves to the abstract conceptualisation stage where they draw some conclusions and plan how to adapt their action in a next iteration through the cycle. In addition, Kolb also discusses four learning styles, each of them related to different stages of the cycle. This theory has later been extended by Honey and Mumford (Honey, 1992), stating that learners can be placed along the two axes highlighted in Figure 4.

![Learning Styles by Honey and Mumford](image)

The perception continuum differentiates between people who prefer thinking and reasoning about the theory of new situations before partaking, and people who prefer to gain an intuition about what to do in a certain context. The processing continuum distinguishes between those who prefer to learn from observing others first and those who prefer to figure things out as they go. Based on these axes, four distinct learning styles are identified, each of them learning best during a different stage of the learning cycle. By ensuring that we cycle through all four stages, we guarantee that any of these four types of learners have a chance to grow their knowledge.

3 SURVEY

In order to learn more about the needs of table tennis trainers, we conducted a survey with trainers registered at the Vlaamse Tafeltennisliga (VTTL) and the table tennis division of Sporta, the two largest Flemish table tennis organisations. From the invited trainers, 82 participated and filled in the survey. A major part of the responses (around 40%) came from trainers with more than ten years of experience, and we got a healthy mix of both, experienced trainers leveraging years of expertise as well as less experienced trainers looking forward to introducing new ideas.

Our survey started with a short introduction about the planned research. This was followed by a number of questions divided into distinct sections, including the trainers’ experience, the material they use during training sessions, the way they analyse players, how they communicate with players, their preparations of training sessions as well as general feedback. The detailed questions that were used in our survey as well as the anonymised results are available online. Based on our literature study and the results of our survey, we derived the following eleven core requirements for a technology-enhanced personalised learning environment for the domain of table tennis:

**R1: Trainers Should Be Able to Manage Exercises.**

From the open questions of our survey, it became clear that trainers would like to have the freedom to experiment with their own exercises and content. This indicates that trainers should be able to inspect as well as add new exercises to a learning environment. For every exercise it should be possible to inspect its name, a description and image, as well as the techniques it teaches and any prerequisites in the form of techniques a user already must know.

**R2: Trainers Should Be Able to Manage the Evaluation Criteria for a Technique.**

The evaluation criterion forms the cornerstone in the feedback process to players (one of the important aspects of Kolb’s experiential learning cycle) and we should strive to improve upon the criteria as well as offer trainers the flexibility to emphasise certain points. In order to achieve this goal, trainers should be able to inspect all the techniques and the way they are related. Further, it should be possible to inspect a technique’s evaluation criteria as well as to add or delete an evaluation criterion.

**R3: Trainers Should Be Able to Manage Assessments.**

Assessments are used to validate whether a player masters a certain technique; a necessary feature if we want our solution to be able to offer only those exercises that are in the zone of proximal development. To ensure that trainers do not feel forced into one way of doing things—which was a major point of feedback in our survey’s open question—assessments should be manageable by trainers. They should be able to add new assessments, declare the outcomes of passing and assessment and select the correspond---
ing exercises (together with the minimal requirements that a player should meet) for an assessment.

**R4: Trainers Should Be Able to Supervise Assessments.** When a player is performing an assessment, a trainer should be able to grade the player’s performance. Based on this grading, the system should be able to update a player’s profile. The reported performance can be used in future communication with a player and improve the communication process, a demand that was indicated by half of the trainers in our survey. Further, the updated user profile allows the system to later filter out exercises that are either too difficult or too easy.

**R5: Trainers Should Be Able to Manage the Player Knowledge Graph.** It should be possible for a trainer to inspect a player’s current knowledge, including their proficiency level for individual techniques. Further, a trainer should be able to manually update the data about a player’s knowledge. This functional requirement enables trainers to fine tune the model in case something went wrong, or to bootstrap the process for newly joining experienced players.

**R6: Trainers Should Be Able to Inspect a player’s Past Performance.** In order to provide proper long-term supervision, trainers need to have a good mental model of the strengths and weaknesses of their players. However, 68% of the survey participants indicated that they do not keep any notes about their players and solely rely on their memory. To better assist trainers, they should be able to inspect the results of a player’s past assessments and training sessions, including the grading of specific evaluation criteria. This is particularly useful in clubs where multiple trainers are teaching the same players.

**R7: Trainers Should Be Able to Prepare Training Sessions.** It should be possible for a trainer to prepare a session with minimal effort, given that in our survey a lack of time was the most common reason why trainers do not prepare sessions. This preparation should include the selection of exercises from the list of recommended exercises in a player’s zone of proximal development (automatically generated based on their current proficiency level for different techniques). A trainer should further be able to select certain exercises based on a specific technique.

**R8: Trainers Should Be Able to Supervise Training Sessions.** While supervising sessions, a tool should first show the exercises to players before they start performing them. Once the specified time for an exercise has elapsed, this should be indicated to the trainer. A trainer should then have the possibility to grade the player’s performance for each of the evaluation criteria of the techniques forming part of the exercise. This also provides an ideal opportunity for the trainer to immediately provide some feedback (based on the evaluation criteria questions) to the player who can then reflect before starting a next cycle as proposed in Kolb’s experiential learning cycle.

**R9: The Application Should Adapt to Different Screen Sizes.** The use of large and heavy devices, such as laptops, can be cumbersome during training sessions. More than half of the survey participants indicated to use their smartphone during a session, while less than 10% use a tablet or laptop. Therefore, an application should adapt to arbitrary screen sizes.

**R10: The Application Should Allow the Sharing of Data between Trainers of a Club.** Currently, one of the main challenges is that the sharing of information between trainers is an ad-hoc process with less than 20% of the survey participants having a system to share information. An application should therefore enable data sharing between trainers of the same club.

**R11: Data of a Club Should Be Private to That Club** The training progress and contact details of members is sensitive information. It should therefore not be possible for members of a different club to gain access to this information.

## 4 LEARNING ENVIRONMENT

Our solution for a personalised learning environment consists of two major parts. First, there is the general conceptual model and framework to represent all the data and metadata necessary for the general knowledge graph as well as a player’s expertise about parts of this graph (player knowledge graph). Second, we have developed a custom progressive web application tailored towards the needs of trainers (Malaise, 2021).

In order to model the complex domain-specific knowledge, we opted to use the resource-selector-link (RSL) hypermedia metamodel (Signer and Norrie, 2007) shown in Figure 5. A Resource is representing any real or virtual object, such as an image, a

---

**Figure 5: Overview RSL hypermedia metamodel.**
student, a specific skill level or an arbitrary document. Often one does not want to refer to an entire resource, but to address parts of it (e.g. a specific person in an image) via the concept of a Selector. Finally, a Link can be used to represent a bidirectional relationship between two or more entities.

These three core types are all subtypes of the more general Entity type and as such can have their own properties. This further implies that a link can not only be used to define relationships between resources but also have selectors or even other links as source or target. It is out of the scope of this paper to describe the full RSL metamodel but further details can be found in (Signer and Norrie, 2007).

The conceptual model of our personalised learning environment for the domain of table tennis is shown in Figure 6. The rectangles represent specific RSL Resources and the directed arrows represent RSL Links between Entities. Further, when a link itself is used as the source of another link, this is indicated via a black solid circle on the source link.

A Player represents a person who is a member of a club for which a trainer is creating training schedules. The goal of our solution is to assist trainers in improving the skills of players and therefore players form a central component of our conceptual data model. A Technique is a certain skill or a piece of knowledge that the player needs to master in their path to proficiency in a sport discipline. A technique might require a certain level of proficiency of other techniques before it can be acquired, which is represented by the Requires link and the MinimalLevel link over this Requires link. For example, a technique could be the act of hitting the ball. Before a player can serve, they should be able to hit the ball. This implies that if a player is having trouble serving, it could be caused by the fact that they have not yet sufficiently mastered the skill of hitting the ball. The ProficiencyLevel is the level at which a player has mastered a specific skill or technique and is based on the three-stage model of motor skill acquisition introduced in Section 2.

The EvaluationCriterion consists of a basic question about parts of a technique (e.g. “Did the player place their feet the correct way?”) and is associated with the corresponding technique via the ValidatedBy link. These questions can be used to evaluate a player’s performance (score between 1 and 5) after each exercise. Further, a trainer might discuss the criteria with the player and provide immediate feedback, pushing the player to reflect and improve. An EvaluationCriterionResult contains information how well the player performed for a specific evaluation criterion. The results for a given evaluation criterion are linked to the exercise performed when the trainer graded the criterion via the PerformedExercise link. This contextual information is useful since when a player meets certain requirements of some exercises but others not, it might indicate underlying issues with other techniques used in the exercises.

An Exercise describes a situation in which a player can learn about or improve on one or multiple techniques as indicated by the Teaches link. Via the Requires link it is possible to indicate that an exercise depends on certain techniques. Before a player
is able to perform the exercise, they need to master a technique at least on the proficiency level indicated by the MinimalLevel link defined over the association between an exercise and a technique. This is necessary to filter out those exercises that are currently too difficult for the learner, in order to stay within their zone of proximal development.

An Assessment is a special set of exercises referred to via a HasExercise link and designed to gauge the proficiency of a player in certain techniques. The HasExercise link further forms the source of the MinimallyRequires link that points to an AssessmentRequirement stating the objective and measurable requirements for a specific exercise in the assessment. It consists of a certain condition that needs to be fulfilled for an exercise in order to say that the player mastered the technique at a sufficiently high level. For example, in an assessment on the ability of players to perform the forehand stroke, there could be an exercise stating that they would need to hit a piece of paper placed on the table with the ball. The requirement could state that within 5 minutes a player should place at least 20 strokes in the indicated area. An AssessmentResult manages information on how a player performed for a given assessment by linking to the corresponding RequirementResults. When an assessment has been passed successfully, it will lead to the creation of new Masters links associating a player with the techniques they now master at a specific level. The level of mastery that has been achieved is represented via the AtLevel link defined over the Masters link. The LearningOutcome that is related to an easement via the Awards link defines which Masters and AtLevel links will be created. Based on a player’s achieved proficiency levels, our application is able to suggest or hide certain exercises.

A Training defines the training schedule for an individual session on a specific day, including all the exercises that need to be performed as indicated via the IncludesExercise link (note that this link will normally point to multiple targets representing the exercises). Information about how the player performed during a training is further managed in a TrainingResult. A training is linked to its result via a HasResult link and training results can provide valuable insights on how a player is evolving and hopefully improving over multiple training sessions.

Aside from general comments about a player’s performance during a training, a TrainingResult normally consists of multiple EvaluationCriterionResults as represented via the Includes link. Thereby, an EvaluationCriterionResult holds the score given by the trainer on how well the player scored on a certain EvaluationCriterion (referred to via the BasedOn link) while performing a specific Exercise (indicated by the PerformedExercise link). This fine-grained model allows us to look for specific patterns or underlying reasons why players score well for an EvaluationCriterion in one exercise while they might perform poorly for the same criterion in another exercise.

While the conceptual data model forms the core of our application, the manual manipulation of the graph via queries would lead to a poor user experience and be too technical for the average trainer. Therefore, we developed a server based on the Spring Boot framework, providing endpoints to request all important derived results as well as to register new information. To interact with the server, a progressive web application that can either be used as a desktop application or on a mobile device has been realised. The web application was developed using the angular framework. The application allows trainers to design personalised training schedules for players and supports the supervision of assessments without requiring a user to know anything about the underlying knowledge graph.

5 USE CASE

We discuss some of the most important aspects of the user facing application as well as how a trainer is going to interact with them. In order to make sure that a trainer can utilise their own expertise and creativity, they have to be able to add their exercises to the system (requirement R1). When doing so, they can specify a name, a duration, an image, any number of tags and a description. They can indicate which techniques need to be mastered at which proficiency level before a player can perform the exercise, and the skills taught by the exercise. This metadata is going to be used to suggest exercises to individual players.

The knowledge representation of a specific player consists of a detailed representation of their achieved proficiencies and knowledge (player knowledge graph). The direct manipulation of this graph would be too technical and we provide trainers the possibility to create an assessment (requirement R3), where they can specify which skills they would like to assess as well as the proficiency level. The application then suggests exercises that are a good fit for the assessment. For every selected exercise, the trainer has to provide a minimal requirement condition needed for a player to pass that part of the assessment, which is later used to modify the player knowledge graph.

---

6https://spring.io/projects/spring-boot
7https://angular.io
8https://youtu.be/OSw2PWpG6dg
When a trainer feels that a player is ready, they can ask them to take part in an assessment (requirement R4). Each exercise is shown with a built-in countdown timer and after time is up, a dialogue asks the trainer whether the player passed the requirement for the exercise. If a player has passed all the requirements at the end of the assessment, they will be awarded the configured proficiencies. In case they failed some exercises, the system will check whether the failed exercises contained some dependencies to skills that did not show up in the dependency list of the exercises the player passed. If such a skill exists, it might likely form a knowledge/skill gap and the system will report this information to the player.

When planning a session for a specific player, our solution will automatically filter out all exercises that are either too difficult or too easy for the player based on the proficiency data stored in their knowledge model. This guarantees that all exercises selected by the trainer will be situated in the player’s zone of proximal development. Our application further allows trainers to search for specific exercises based on a name, tags or specific techniques that they would like the player to work on (requirement R7).

When supervising a training session, the trainer is shown a screen with the image, name and description of the exercise as well as some controls to operate the system. Once the time for an individual exercise has elapsed, a list of evaluation criteria for the techniques taught by the exercise, as well as the techniques directly required by these techniques, are shown to the trainer for detecting errors in prior techniques as well as knowledge and skill gaps. Further, the trainer can immediately provide this detailed feedback to a player, supporting the reflective observation and abstract conceptualisation phase of Kolb’s experiential learning cycle (requirement R8). Once the training session is completed, a trainer can add more general feedback for future analysis (requirement R6).

Our solution contains some dedicated views to inspect the progress of individual players (requirement R6), including a spider chart of the most recent average evaluation criteria scores of techniques they have been working on, a graph showing the evolution of their scores for certain techniques over time, as well as an enhanced knowledge graph view in which a trainer can see and modify the achieved proficiency of each technique (requirement R5) in the form of a player knowledge graph as illustrated in Figure 7.

While our solution is well suited for one-to-one sessions, various trainers asked us to also include some functionality to support group sessions. In order to address the needs of these trainers, we decided to also add a group mode where an exercise will only be available if each group member satisfies the necessary requirements. Further, there is a single evaluation window, containing a table with all the techniques used throughout the training on one axis and all participating players on the other axis, where a trainer can report a player’s performance for a specific technique (requirement R8).

6 EVALUATION

The feedback from an experienced group of trainers is essential to ensure that our application with its knowledge graph-based approach will not only be useful in real life, but that trainers would also like to use the application. While various table tennis clubs indicated to be available and willing to participate in trial runs, it was unfortunately not yet possible to perform an in situ evaluation due to the global pandemic after the Covid-19 outbreak. The best alternative was the creation of a video clip with some voice-over describing the application and illustrating how the tool could be used by a trainer, both during the preparation and the supervision of a training session. The video was then emailed to all the trainers who indicated to being open for further questions in our initial survey, and we received a total of 15 responses from trainers filling in the second survey. Everybody filling out the survey was invited for a 30 minute video interview to discuss our solution and provide further feedback, and four trainers accepted to participate in an interview.

When asked whether they would like to use our solution in their club, three out of the 15 participants answered with a neutral score of 3 on a 5-point Likert scale, six trainers answered with a score of 4 and another six participants answered with a score of 5. These scores indicate that most trainers seem to be excited about the idea of using the tool and the results are quite promising. We noticed some recurring feature requests such as providing an offline first version and improving the mobile user experience.
The trainers who took part in the interview all showed enthusiasm and were excited to use the system in a real environment. A main topic in most interviews was the question of who should be able to modify the general knowledge graph, as well as whether we should push for exercises created in one club to automatically be available to all clubs; an interesting topic that will need some further investigation.

7 DISCUSSION

We presented a prototype of a personalised learning environment helping table tennis trainers in preparing tailor-made training sessions. Our conceptual model is not limited to table tennis, but could already be used as is in tools for most individual sports disciplines. Further, our conceptual model might be generalised to support more traditional educational settings as well as hybrid classroom setups where on-site classes are combined with partial self-study trough e-learning. For instance, the three-stage model of motor skill acquisition might be replaced with Bloom’s taxonomy (Bloom, 1956). Further, assessments and the analysis of skill gaps should translate well to remote e-learning environments and private tutoring sessions.

In addition to suggesting the right content for learners, our model could also be used to efficiently analyse the current knowledge levels of newcomers. A student might be given a set of exercises about a specific skill and if they fail to perform the exercise adequately, we can suggest exercises based on the direct requirements of that skill. If a student manages to complete that exercise, it implies that the prerequisite skill is not the issue; otherwise we need to analyse where the knowledge gap causing the student to fail that prerequisite is coming from.

Past research has shown that students perform better when the material has been adapted to their preferred learning styles (Mustafa and Sharif, 2011). A main advantage of our model being based on the RSL hypermedia metamodel is that it allows us to take personalisation even further. Instead of simply suggesting different exercises based on a learner’s proficiency, we might adapt the exercises based on the RSL model’s concept of structural links and their use for adaptive document structures (Signer, 2010).

8 CONCLUSION

We have presented a prototype of a technology-enhanced personalised learning environment for the domain of table tennis, making use of knowledge graphs in combination with the results of assessments to suggest exercises in a player’s zone of proximal development. The discussed research on a conceptual model and domain-specific application represents a step towards a personalised learning environment where the learner is central. We do not only aim to provide the right content at the right time, but also envision to further adapt the presented content based on the underlying RSL hypermedia metamodel in combination with a learner’s preferences, their previous experience as well as their learning style.

REFERENCES