So, H. J. et al. (Eds.) (2020). Proceedings of the 28th International Conference on Computers in Education. Australia: Asia-Pacific Society for Computers in Education

# Learner Model of Knowledge Grounding in Discovery Learning

**Takuya WADA[a*], Yuki HAYASHI[a], Kazuhisa SETA[a] & Tomoko KOJIRI[b]**
[a] *Graduate School of Humanities and Sustainable System Sciences, Osaka Prefecture University, Japan*
[b] *Faculty of Engineering Science, Kansai University, Japan*
* ddb04005@edu.osakafu-u.ac.jp

**Abstract:** Discovery learning is a learning method for acquiring new knowledge and utilizing such knowledge in the real world. As a framework for support discovery learning, several exploratory learning environments (ELEs) have been developed, which aim to acquire and utilize knowledge within such environments. However, some learners cannot utilize their knowledge because it is weakly linked to entities in the real world. Few ELEs aim at using learning to strengthen the connection of the learner's knowledge with real-world entities. In view of the insufficient framework for supporting a discovery learning that links knowledge with entities, this study identifies the cognitive state of errors in the grounding of discovery learning, and internal representations of information necessary to detect such errors in an ELE.

**Keywords:** discovery learning, learner model, exploratory learning environment, grounding, biology learning

## 1. Introduction

In recent years, various types of active learning have been introduced in primary and secondary school education, among which discovery learning is one of them. The purpose of discovery learning is to acquire new knowledge and to utilize such knowledge in the real world. Choi et al. (1995) stated that some learners are unable to utilize such knowledge in everyday circumstances, and that activities utilizing knowledge in the real world should be emphasized. Knowledge in the field of natural science mainly refers to matter or phenomena in the natural world. If a learner's knowledge is not connected (hereinafter called "grounded") with matter or phenomena in the natural world as entities, the learner may not be able to utilize it. For learners to be able to utilize knowledge in the real world through discovery learning, a type of learning that strengthens the learner's grounding with real-world entities is necessary. This study considers knowledge as concepts and their relationships, and entities as external resources that compose the real world. Grounding is regarded as a state in which the knowledge of the learner is connected with the correct entities.

De Jong et. al. (1998) stated that learners should have the knowledge beforehand if discovery learning is to be fruitful. Utilization of this knowledge is a prerequisite for improving the quality of the learning outcomes. However, applying knowledge to the learning targets requires their appropriate grounding. If the grounding is incorrect, the knowledge cannot be used effectively. For example, to identify the species of observed unknown plants, the learner needs to utilize knowledge of characteristic plant parts to identify them. However, if the learner cannot ground plant parts of the knowledge with parts of the unknown plant, the knowledge is not utilized well. As a framework to support discovery learning, several exploratory learning environments (ELEs) have been developed. Santos et al. (2012) stated that the interaction of learners is highly unstructured and is difficult to model and prescribe. They proposed an approach to develop intelligent support in ELEs. However, a specific way to identify the status of a learner's grounding and types of log data as useful information for the identification has not been proposed.

Because a general method for evaluating whether a learner can apply knowledge to an entity is yet to be clarified, the cognitive state that indicates the accuracy of the grounding is unknown. Therefore, this study deals with the question of what mechanism can capture a cognitive state through which knowledge can be correctly applied in the real world. This study identifies a cognitive state of errors in the grounding of discovery learning, as well as internal representations to detect such errors from interactions between an ELE and a learner.

## 2. Related Studies

Conventional ELEs promote the utilization of knowledge by providing a virtual learning environment that does not cause grounding errors. Table 1 shows characteristics of each type of ELE. Simulators and microworlds are virtual worlds generated on a computer with the aim of encouraging learners to discover specific knowledge through direct operations and observations. Simulators offer a virtual world as a representation of a domain model. A flight simulator is an example of a simulator, as indicated by Perry (2004). Microworlds represent the domain model and allow learners to add or remove elements. As an example of a microworld for physics learning, a virtual environment in which learners can change the conditions of a moving object and its surroundings and observe their behavior, imitating the real-world, was developed by Cockburn et al. (1995). In general, simulators and microworlds cannot assess a learner and do not have a diagnostic function for intelligent support.

An intelligent microworld is a learning environment using a learner model and intelligent support functions to help learners who are struggling with using a microworld. There are several learner modeling techniques applied in an intelligent microworld. An intelligent microworld for chemical experiments developed by Yoshikawa et al. (2000) records the operational sequence of the learners and estimates their goals as well as the errors in their operating procedures and knowledge. In addition, a physics-based intelligent microworld developed by Reid et al. (2003) estimates laws that the learner is unaware of as based on experimental results and the derived laws inputted by the learner.

An error-based simulation (EBS), which is a method for generating phenomena by applying erroneous concepts used by students, is a promising method allowing the students to be aware of errors based on certain correctly known phenomena connected to their proper concepts. EBS promotes the correction of learner errors in knowledge by providing an unreal world that expresses the learner's erroneous knowledge as entities. Because the learning target is a correction of specific knowledge errors, EBS is introduced in a learning context with explicit tasks. Hirashima et al. (2009) defined the drawing of a force direction based on errors in the learner’s knowledge, and their developed EBS system estimates errors in the learner’s knowledge from drawings inputted by the learner.

Table 1

*Characteristics of several types of ELE*

| | | ELE | | | |
|---|---|---|---|---|---|
| | | simulator | microworld | intelligent microworld | EBS |
| characteristic | learning task | given | learner decided | learner decided | given |
| | intelligent support | without | without | with | with |
| | learning target | knowledge acquisition | knowledge acquisition | knowledge acquisition | knowledge acquisition |

Most ELEs target the acquisition of knowledge, and there are few ELEs that target the correct grounding of the learner. Although some ELEs support error correction of knowledge to reach the correct grounding such as EBS developed by Hiramoto et al. (2018), few models exist that capture cognitive states by distinguishing knowledge errors from grounding errors. In such ELEs, although the learner can temporarily apply their knowledge to only entities provided by ELEs, they may not be able to apply their knowledge correctly in the real world, which allows grounding errors to occur. For example, suppose that the learner applies knowledge of angiosperm to an entity of a plant that the ELE shows as an angiosperm. Then, even if the learner observes other angiosperms in the real world, the learner may not be able to apply their knowledge of an angiosperm to these entities. In this case, the correct application of knowledge requires having knowledge regarding the characteristics of the angiosperms and a correct grounding of such knowledge and the characteristics of the entities. The learning target focused on in this study is a grounding of knowledge and entities in virtual worlds on computers, and a learner modeling method for detecting grounding errors is established. This study focuses on complex cognitive states in a learning context in which the learner can decide on a learning task and observe numerous targets, identifying the cognitive states of such errors.

## 3. Types of Errors of Grounding and Representation Method

### *3.1 Errors of Grounding*

This study regards knowledge as a binary relation between two concepts and focuses on hierarchical knowledge structures as complex knowledge structures. The target knowledge structure in this study consists of "is-a" relationships, which indicate inheritance relationships between concepts, and "has" relationships, which indicate ownership relationships between concepts, and covers the domains of the plant types and their parts. Because discovery learning for many students is carried out after learning each concept of the domain as prior knowledge, learners who have formed at least individual concepts of a domain knowledge structure correctly are targeted. Although these learners have individual concepts, the relationships between their concepts may be incorrect.

Correct grounding is a state in which each knowledge concept is applied to the correct entities, and the correct relationship of the knowledge is applied to the relationship between entities. Errors of grounding during learning with errors of knowledge and grounding are as follows:

I. The relationship between concepts of a knowledge structure is incorrect.
II. The combination of the applied concept of a knowledge structure and an entity is incorrect.
III. The relationship between entities captured by the learner is incorrect.

Error I is a state in which the learner cannot properly apply knowledge to the correct entity because the knowledge is incorrect. For example, a learner who mistakenly believes that a "pine tree has coconuts" cannot be grounded correctly because the learner has a discrepancy in the observation result that "pine trees do not have coconuts" when observing the pine tree as an entity. Error II is a state in which individual concepts are applied to incorrect entities. For example, the learner recognizes a palm entity as a pine tree or recognizes a coconut entity as a conifer cone. Error III is a situation in which the learner overlooks the entity and misunderstands the relationship between different entities. For example, the learner recognizes that "the palms do not have coconuts" when observing a palm entity and overlooks the coconut. In this case, even though the learner has the correct knowledge that "a palm has coconuts," the learner cannot ground the knowledge correctly owing to a discrepancy between the ownership relationships of the knowledge and entities. These individual errors occur concurrently. To define the state of the learner in which these errors occur, this study proposes a representation method of the cognitive state of the grounding of knowledge and entities. Furthermore, the state of the learner defined based on the representation method is formulated in a computer readable manner.

### *3.2 Representation of Errors of Grounding*

As a learner's concept system in discovery learning, we consider concepts acquired during classroom learning and concepts formed through observation. This study calls the former a "knowledge layer" and the latter an "observation layer." Species of plants include individual plants (e.g., pine trees and palms) and groups include multiple plants (e.g., seed plants and gymnosperms). There are inclusive relationships among them. Species of plants have parts in common (e.g., stamens and calyxes), and there are ownership relationships between species of plants and parts of plants.

Knowledge and observation layers can be represented as a concept map of plants and their constituent parts. In the concept map, two types of nodes representing "species" and "parts" of plants are introduced, and the inclusion relationship of the plant types is represented through an "is-a" relationship link, whereas the ownership relationship between concepts is represented as a "has" relationship link. A specific concept of species inherits the "has" relationships of its generic concepts. In addition, "not has" relationships are introduced in the observation layer to explicitly express a cognitive state in which the learner is aware of a nonexistence of a part of an observed plant. This study approximates a state in which relationships between a concept of knowledge layer and an entity of observation layer and relationships between relationships in both layers are correct as correct grounding. These relationships are represented as “instance-of” relationship links. The correspondence between links of both layers is also represented as an “instance-of” relationship link because relationships between concepts are also observed as relationships between entities.

Figure 1 shows an example of the proposed concept map. The entity connected to the concept “XXX” by the “instance-of link” is described as “#XXX”. In the knowledge layer, is-a links indicate that “pine trees and palms are gymnosperms,” and “has links” indicate that “gymnosperms, pine trees, and palms have male flowers,” “pine trees have pineapples,” and “palms have coconuts.” The pine trees and palms in the knowledge layer inherit the “has link” of male flowers from the gymnosperms. In the observation layer, “a pine tree has male flowers” and “a pine tree does not have coconuts” are recognized through observation. Regarding an “instance-of” relationship between the relationships of two layers, which is omitted in Figure 1, the correspondence of the “has link” between “pine trees” and “male flowers” in the knowledge layer and the “has link” between the “#pine tree” and the “#male flower” in the observation layer indicates the correct instance-of relationship. By contrast, because the “has link” between the “pine trees” and “pineapples” in the knowledge layer does not correspond with the “not has link” between the “#pine tree” and the “#pineapple” in the observation layer, this instance-of relationship between the relationships of both layers is incorrect.

Error I corresponds to an error of a link in the knowledge layer. Error II corresponds to an “instance-of link” between both layers. Error III corresponds to an error of a link in the observation layer. The types of errors in these relationships are as follows:

a). The “has link” in the knowledge layer is incorrect (Error I).
b). The “is-a link” in the knowledge layer is incorrect (Error I).
c). The “instance-of link” is incorrect (Error II).
d). The “has / not has link” in the observation layer is incorrect (Error III).

In the binary relationship between the knowledge layer and the observation layer, the knowledge and entities are grounded correctly when “instance-of links” between concepts and the “instance-of link” between “has links” are correct. A state in which these “has links” and the “instance-of link” are correct indicates that the “instance-of” relationship between the relationships of two layers is also correct. Therefore, an error in grounding is exposed as an error of the “has link” in the knowledge layer, an error of the “instance-of link” between concepts, or an error of the “has / not has link” in the observation layer. Although an error of the “is-a link” is not exposed as a grounding error, the “has link” incorrectly inherited by the “is-a link” is exposed as such an error.

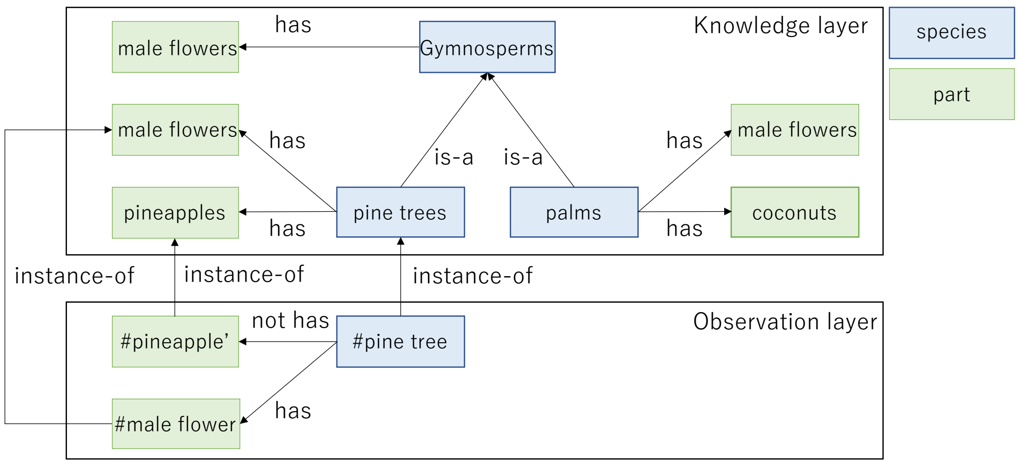


*Figure 1.* Example of the proposed concept map

## 4. Learner Modeling of Grounding State

This study focuses on a discovery learning in which the domain has a knowledge structure consisting of inclusion relationships and ownership relationships between concepts, and proposes a learner model that expresses grounding errors, as described in Section 3. The domain has a correct knowledge structure with multiple hierarchies of inclusion relationships and single hierarchies of ownership relationships. The correct concept system consists of the knowledge structure of the domain and the corresponding entities are defined. To consider "instance-of links" between the learner's concepts and "instance-of links" between "has links," the learner's partial grounding state is expressed by understanding the state of the binary relation of the knowledge layer and the corresponding binary relation of the observation layer of the correct concept system.

Figure 2 shows a part of the correct concept system corresponding to the learner's partial grounding state. "species X" is any concept in a correct concept system, and the part that "species X" has is described as "part x." In addition, "#species X" is described as the correct entity of "species X," and "#part x" is described as the correct entity of "part x." By integrating the partial grounding states, the learner's grounding states for the overall correct knowledge system is modeled.

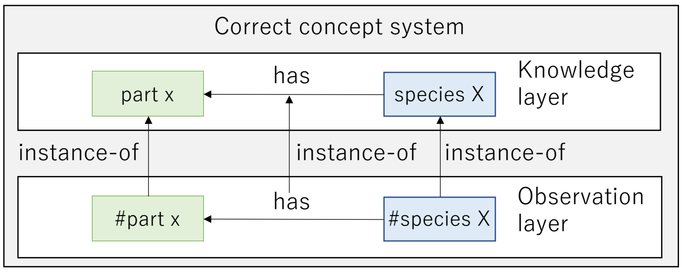


*Figure 2.* Part of the correct concept system corresponding to the learner's grounding state

Discuss internal representations to detect errors in Section 3. Table 2 shows a learner model that summarizes attributes and values that the computer must capture to detect each error.

Table 2

Internal representations to detect errors

| attribute | value |
|---|---|
| a1. Existence of "has link" between "species X" and "part x" in the knowledge layer | true, false |
| a2. Existence of "#species X" in the observation layer | true, false |
| a3. Existence of "#part x" in the observation layer | true, false |
| a4. State of "has link" of "#part x" | has, not has, null, with_other |
| a5. State of "instance-of link" of "#species X" | true, false, null |
| a6. State of "instance-of link" of "part x" | true, false, null |
| a7. State of "instance-of link" of "#species X" | true, false, null |
| a8. State of "instance-of link" of "#part x" | true, false, null |

To detect error a, it is necessary to obtain the information about the "has links" in the learner's knowledge layer. Error a is a state in which the "has link" of the knowledge layer of the correct concept system is not in the learner's knowledge layer. In addition, the state in which the "has links" that are not in the correct concept system exists in the learner's knowledge layer is also error a. Therefore, the learner model requires the attribute of existence of the "has link" between "species X" and "part x" in the knowledge layer. For the value of this attribute, "exist" is described as "true" and "not exist" is described as "false."

Error c is a state in which there is no "instance-of link" between correct concepts of both layers, and these concepts have an "instance-of link" with the incorrect concept. To detect error c, the learner model requires the attributes indicating the state of the "instance-of links" for each species X, "part x," "#species X," and "#part x." For the value of these attributes, "correct" is described as "true," "incorrect" as "false," and "not existence" as "null." The error of "instance-of link" between the "has links" of two layers is represented by the error of the "has link" in knowledge or observation layer.

To identify error d, information of the existence of "#species X" and "#part x" in the observation layer and "(not) has link" between these concepts is required. Error d is a state in which "#species X" and "#part x" are in the observation layer, and these concepts have a "not has link." The state in which “#part x” in the observation layer has a “has link” with a concept other than “#species X” is also error d. Therefore, to detect such errors, the learner model requires the attributes of existence of “#species X” and “#part x” in the observation layer. For the value of this attribute, “exist” is described as “true,” and “not exist” as “false.” Furthermore, an attribute that indicates the state of “has link” of “#part x” is also required. The value of this attribute is described as “has,” “not has,” “null (i.e., no link), and “with_other (i.e., “#part x” has a “has link” with a concept other than “#species X”).”

The learner's grounding state can be approximated by obtaining values of all attributes in Table 2 for all binary relation of the overall knowledge and observation layer. This modeling assumes that the ELE is designed such that values of all attributes in Table 2 can be obtained from interactions between an ELE and a learner. By updating the learner model every time the learner performs a specific operation, the grounding state of the learner can always be captured.

## 5. Conclusion

This study focused on discovery learning for strengthening the grounding between knowledge and entities in the real world, for which a framework for learner modeling was proposed. Specifically, the types of errors in the grounding of discovery learning were identified, and the information necessary for a computer to detect such errors was suggested. These errors are formalized into computer-readable attributes and values, and the learner model for ELE that can obtain such information was proposed.

Currently, a knowledge structure with single hierarchies of ownership relationships is being targeted. However, as an example, some plant parts also have its more detailed parts in the real world, and therefore, in a further study a learner model for a knowledge structure with multiple hierarchies of ownership relationships should be developed.

In some cases, learners may be able to correct a grounding error themselves. Although the proposed learner model can capture individual errors in the grounding of the learner, this model cannot determine whether a learner is struggling from such grounding errors. The state of errors in the grounding and knowledge indicating the struggles of a learner must be identified, and a learner model that can capture them should be developed.

## Acknowledgements

The work was supported in part by JSPS KAKENHI (No. 19H01725).

## References

Choi, J. I., & Hannafin, M. (1995). Situated cognition and learning environments: Roles, structures, and implications for design. *Educational Technology Research and Development*, *43*(2), 53-69.

Cockburn, A. & Greenberg, S. (1995). TurboTurtle: A collaborative microworld for exploring Newtonian physics, *CSCL Conference on Computer Supported Collaborative Learning, 1*, 62-66.

Gutierrez-Santos, S., Mavrikis, M. & Magoulas, G. D. (2012). A separation of concerns for engineering intelligent support for exploratory learning environments. *Journal of Research and Practice in Information Technology*, *44*(3), 347.

Hiramoto, C., Tomoto, T., Horiguchi, T. & Hrashima, T. (2018). Development of an Error-Based Simulation for Understanding Relationship to force, acceleration, velocity and movement. *Proc. of SIG on Advanced Learning Science and Technology,* 5(3), 79-84.

Hirashima, T., Imai, I., Horiguchi, T., & Toumoto, T. (2009). Error-based simulation to promote awareness of errors in elementary mechanics and its evaluation. *Proc. of Artificial Intelligence in Education.* 409-416.

Perry, A. R. (2004). The Flightgear flight simulator. *Proc. of the USENIX Annual Technical Conference, 686*.

Reid, D. J., Zhang, J. & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning, 19*(1), 9-20.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179-201.

Yoshikawa, K., Takahashi, I. & Itoh, Y. (2000). Generating interactive explanations by using both images and texts for Micro World. *Proc. of International Conference on Computers in Education 2000, 1*, 643-650.