Exploring learning in basic logic using stored Log Data and Exam Data

David Jakobsen[1], Steinar Thorvaldsen[2,3] and Peter Øhrstrøm[1]

1: Department of Communication and Psychology, Aalborg University, 9000 Aalborg, Denmark
{poe,davker}@ikp.aau.dk
2: Department of Education, NLA University College, 5812 Bergen, Norway
3: Department of Education, UiT- the Arctic University of Norway, 9037 Tromsø, Norway
steinar.thorvaldsen@uit.no

Abstract. This paper presents how the analysis of log data generated in the context of a basic logic course may be used in the evaluation of the students' learning during the course. It is shown how the analysis of the log data may give rise to rather detailed and profound knowledge on the learning challenges experienced by the students. In this way, we suggest that the systematic and detailed analysis of log data may lead to a more general approach to the evaluation of the learning during the courses in basic logic. One important observation is that the log data shows to what extent the students are able to benefit from a longer response time and therefore to what extent they can make use of the tools available to them. This may provide an important understanding of the students' problems and challenges when working with basic logic.

Keywords: Log data from courses, learning analytics, learning tools, basic logic, syllogistics, propositional logic, validity of arguments, understanding basic logic.

1 Introduction

A course in basic logic has been offered during 2023 to 2nd year students of communication at X University[1]. The course was taught at the A campus and B campus of X University. The focus of this course is the study of logical validity of arguments formulated in terms of elementary propositional logic and classical syllogistics.

Learning Analytics, as a discipline, is "the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environment in which it occurs" [1]. Learning logic falls under this category as well and it has earlier been demonstrated how the learning during a course of this kind can be evaluated based on the logged answers given by the students regarding the validity of arguments presented to them by the teaching system, Syllog and Proplog (see [3], [4], [5], [6]). For each group of answers the score can be calculated as the fraction of arguments that have been evaluated correctly as valid or invalid. It has been shown that the students' scores are rather low, although it is possible to raise

[1] We have anonymized the name of the university to protect the privacy of the participants in the study.

them a bit during the course. But why are the scores so relatively low, even after a traditional course in basic logic? The main purpose of this paper is to provide at least a partial answer to this question.

It is obvious that some students find the some of the topics and problems they are facing even in basic logic extremely difficult. Clearly, it is essential that the teacher understands the nature of the problems and challenges the students are facing. As we shall see, it appears that this kind of better understanding can be obtained through a systematic and detailed analysis of log data generated during the course. Specifically, we shall be interested in exploring the relation between learning and response time as the 'contextual evidence' [2] needed to shorten the inferential distance between claim and observation concerning the student learning.

In the following we shall discuss the course mentioned above and the use of stored log data generated from the student activities during the course. It will be demonstrated how important insights may be extracted from the analysis of the log data. It will be obvious how the procedure may give rise to a more general approach to the teacher's course evaluation.

2 Description of the course

The course in Logic and Argumentation falls in the 2nd year of the education in communication at X University. The students are often surprised when they are introduced to symbolic logic, truth-tables and methods to prove validity etc. Apparently, they associate it mainly with what they encountered in high-school mathematics. One important ambition during the course is to demonstrate that basic logic is in fact closely related to crucial aspects of human communication in general.

144 students, 89 in A and 54 in B, took the course during the period from 6 September 2023 to 6 January 2024. The lectures and exercises were based on a textbook [7] along with various notes. The course covered the following topics: the history of logic, basic syllogistics and propositional logic, reasoning, formal ontology illustrated using Protégé. During a three-day exam at the end of the course the students had to answer a set of questions in writing individually. A part of that exam consisted in using an exam module to answer 10 unique syllogisms and 10 unique propositional arguments. The exam module makes use of two other modules developed for the course, Syllog and Prolog (presented in the sections below).

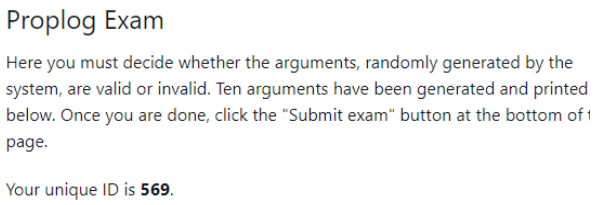


Figure 1: The student is introduced to the exam and gets a unique ID which will be used when correcting the exam.

When working with the exam module, the student will get, 10 propositional (or 10 syllogistic) arguments to evaluate using the tools and methods taught in class. Having

checked an argument for validity, clicking ‘valid’ or ‘invalid’, the student proceeds to the next and after 10 arguments ‘submit’ must be pressed.

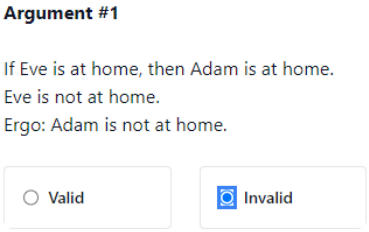


Figure 2: The argument shown here is invalid since the conclusion does not follow from the premises. After evaluating the argument as valid or invalid the student proceeds to the next argument.

Any use of the exam module as well as any use of Syllog and Proplog during the course is logged anonymously.

3 The use of Syllog and Proplog during the course

The teaching makes use of exercises from the two online tools developed for the teaching, Syllog[2] and Proplog[3]. The tools are constructed as games in which the students, when clicking ‘new argument’ will be presented with a new syllogism or propositional argument. The student must then decide whether the argument is valid or invalid and make use of the tools taught on the class and linked to at the website.

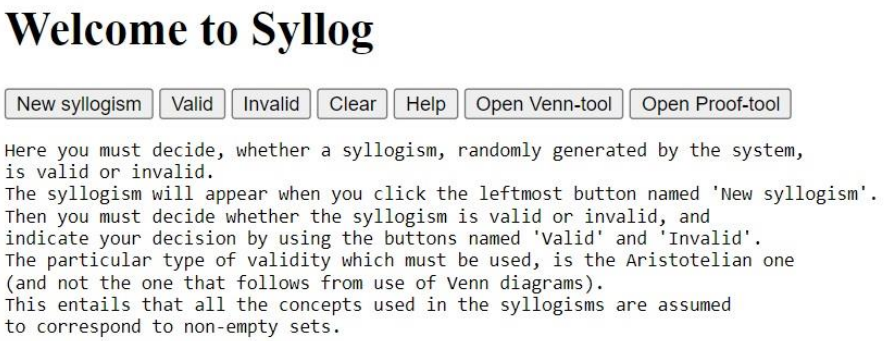


Figure 3 The online tool ‘Syllog’ where the student must decide whether a given argument is ‘valid’ or ‘invalid’.

The students have been taught to evaluate the validity of Syllogisms by 1) transforming the syllogism into canonical form on the basis of which the student can decide whether it is figure 1, 2, 3 or 4. When this is done it is possible, on the basis of lists constructed by medieval logicians using Aristotelean ideas (see [8]), to decide whether it is valid or invalid. If the student discovers that a syllogism is valid, she is taught to prove the validity by the online Proof tool. Here the student inserts the premises in the boxes

[2] https://logic.aau.dk/syllog/webapp/
[3] https://logic.aau.dk/proplog/applet/

which can only be correctly done after the student has transformed the syllogism into logical form. When the premises and the conclusion of the syllogism are inserted, she will be given a list containing the two premises and the negated conclusion. The task is now to derive a contradiction from these three propositions using the three logical rules, TRANS, MUT and EX, which may be presented as follows:

TRANS: $(a(X,Y) \wedge a(Y,Z)) \rightarrow a(X,Z)$
$(a(X,Y) \wedge e(Y,Z)) \rightarrow e(X,Z)$
MUT: $i(X,Y) \rightarrow i(Y,X)$
$e(X,Y) \rightarrow e(Y,X)$
EX: $a(X,Y) \rightarrow i(X,Y)$
$e(X,Y) \rightarrow o(X,Y)$

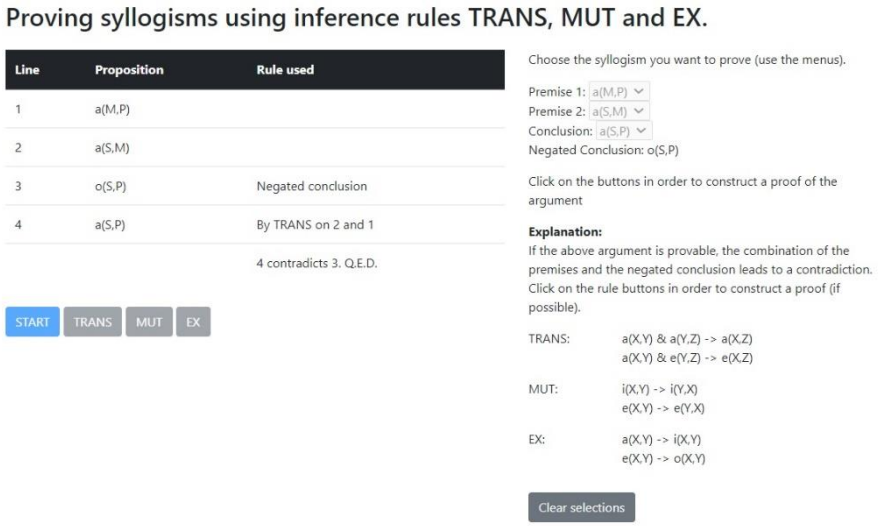


Figure 4: The proof tool will guide the student into proving validity for syllogisms. In this case it is the Barbara syllogism from which a contradiction is derived of the negated conclusion by the use of the rule TRANS on 2 and 1.

In case the syllogism is invalid the student uses the Venn tool where she must demonstrate that it is possible to construct a Venn diagram setting (a model) in which the premises are true, but the conclusion is false.

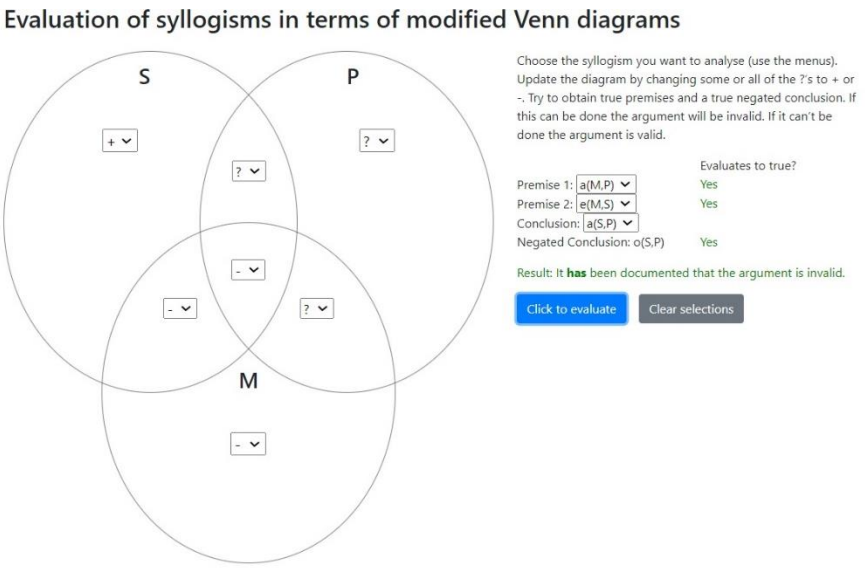


Figure 5: A Venn diagram filled out to make premise 1 a(M,P) and 2 e(M,S) true, and also the negation of the conclusion true o(S,P) which demonstrates that the argument is invalid.

The students also work with Proplog. As with Syllog, they must decide on valid/invalid of the arguments presented, having some tools available from the course:

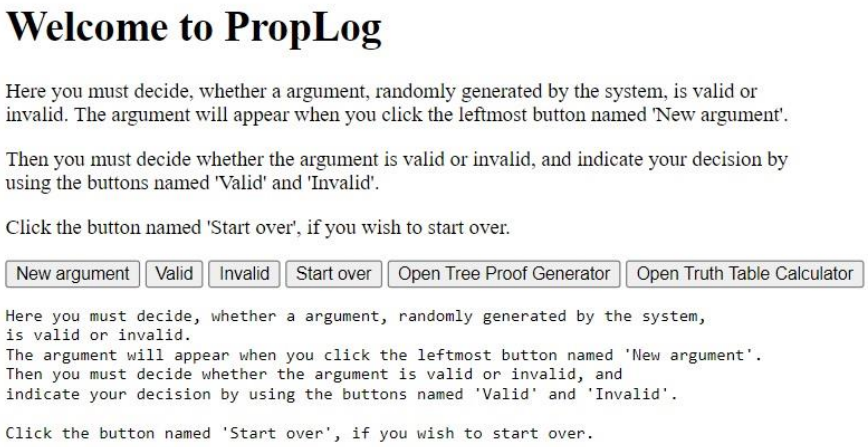


Figure 6: The student clicks on 'New argument' and may then use the available tools.

The process of checking validity/invalidity of a propositional argument is quite different from what is done in case of syllogisms. Working with a propositional argument, the student will first have to transform it into logical form. Consider, for instance,

If Adam is at home, then Eve is not at home
Adam is at home
Therefore: Eve is not at home

This argument can be formalized using p (Adam is at home), and q (Eve is not at home), and put into the form: $p \rightarrow q$, $p \models q$. Following this transformation, a standard truth-table may be constructed to determine the validity of the argument. The student can here take various strategies. She can go all the way and demonstrate that it is a tautology, i.e., that the sequent $((p \rightarrow q \wedge p) \vdash q)$ is a tautology, or she can simply demonstrate in a truth-table that every time the premises are true the conclusion is also true. The general impression is that students are better at proving validity/invalidity for syllogisms than for classical propositional logic.

4 Log data stored during the course

The system automatically logs the data of the students' use of Proplog and Syllog which also includes the exam module for both as well. Various data is logged. For Syllog we log the date and time and the time-lapse of the various answers given to the syllogisms (i.e. the response times). We also log the correct/incorrect score, whether the conclusions of the given syllogisms are true or false and finally we log the figure (1-4), quantor (a,e,i,o), and term (S,M,P) of the syllogism.

For Proplog the data logged concerns date and time and time-lapse, the validity value (valid/invalid) of the argument, which number of the 32 included arguments it was and finally what rule was used to solve the argument (modus ponens, modus tollens, disjunction, not listed). The data is used for two overall goals. The first is to do research and improve the teaching based on research and also to make it possible to evaluate the student's progress during the course.

The results based on the log data from the whole course and exam period can be summarized in the following manner:

Table 1: The results based on log data from the whole course period.

	Correct	Incorrect	Score
Syllog	1963	1242	0,61
Proplog	2468	2360	0,51

The data support strong statistical evidence against the presumption that tasks on syllogistics and propositional logics will be handled equally well (p-value < 0.001 by the $\boldsymbol{X}^2$-test). As it has been observed in our earlier studies, the Syllog-score is higher than the Proplog-score. However, both scores are rather low. But why is this so? What is behind the numbers?

5 An analysis of students' challenges in syllogistic logic

Again, we take all the log data together for the whole course period.

Table 2: General results based on log data from the use of Syllog.

Syllogisms	Correct	Incorrect	Score
Valid	1204	360	0,77
Invalid	755	886	0,46

The results show that the valid syllogisms are significantly easier to detect as such than the invalid syllogisms (p-value < 0.001 by the $\boldsymbol{X}^2$-test).

The development in score during the first three minutes is shown in the following table and diagram:

Table 3: The Syllog score relative to the response time.

Response time	Answers	Score
0-2	1202	0,49
2-20	867	0,59
20-60	429	0,72
60-120	223	0,76
120-180	129	0,78

As pointed out in [2] response-time is a useful context evidence for cleaning data, which is also the case for getting a firmer grip on the log data for the answers. An obvious interpretation of the first short period in this table is that most of the answers given within the first two seconds are based on simple guessing and not very much reflection and reasoning on the aspects and meaning of the arguments in question. As expected, based on how Syllog is constructed the score is close to 0,50.

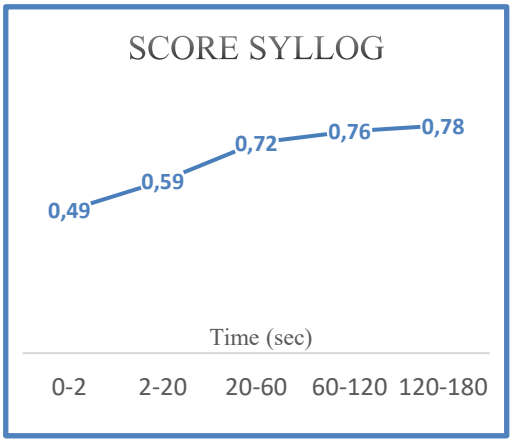


Figure 7: The Syllog score relative to the response time.

After the first two seconds the score is increasing during the whole period with a slightly decreasing rate. This means that the users benefit from having more response time. It seems very likely that the users during the longer response time can make use of the tools (syllogistic proofs and Venn diagrams) that have been introduced during the course.

6 An analysis of students' challenges in propositional logic

A similar study can be carried out for propositional arguments:

Table 4: General results based on log data from the use of Proplog.

Propositional arguments	Correct	Incorrect	Score
Valid	1856	619	0,75
Invalid	612	1741	0,26

It seems that in this case the invalid arguments are even harder to detect compared with valid ones (p-value < 0.001 by the X^2-test). In this case the difference is much more remarkable than in the case of syllogistics. It must be admitted that the score of the invalid propositional arguments (0,26) is surprisingly low. It would be interesting to learn more about reasons behind this.

The Proplog-system is designed with 16 valid and 16 invalid schemes of propositional argumentation. However, there is none of the 16 invalid propositional arguments which appear to give results that differ a lot from the others. It seems that

the problem related to the invalid propositional arguments is general. The question the teachers should ask here is why so many (in fact a majority) of the students think that an argument like: $q \rightarrow \sim p, \sim p \models q$, is valid. In other words, why are so many students ready to accept the following as valid?

If Eve is at home, then Adam is not at home.
Adam is not at home.
Ergo: Eve is at home

Could it be that many students simply don't see the difference between the above argument and the following argument?

If Eve is at home, then Adam is not at home.
Eve is at home.
Ergo: Adam is not at home.

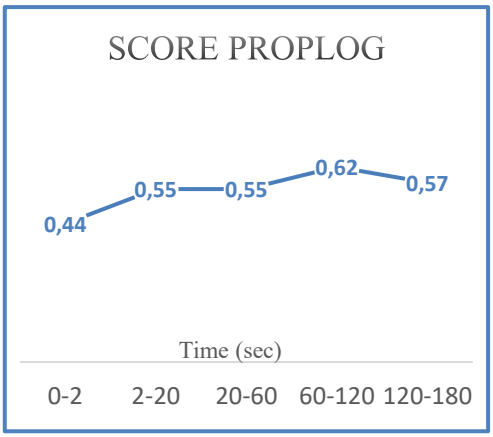
The latter argument was correctly accepted as valid by the vast majority of the students. In fact, the score for the argument, $q \rightarrow \sim p, q \models \sim p$, was 0,88 during the course period.

What is behind this misunderstanding is apparently that it is rather easy in common discourse to jump from "if" to "if and only if", i.e., from a conditional to the corresponding biconditional (see [9]).

As in the Syllog case, it is interesting to look into the log data of the scores relative to the response times. The following table and diagram show how the score varies with the response time:

Table 5: The Proplog score relative to the response time.

Response time	Answers	Score
0-2	2029	0,44
2-20	1840	0,55
20-60	531	0,55
60-120	130	0,62
120-180	60	0,57



The development in score relative to the response time from zero to three minutes is shown in Table 5 and in Figure 8. It is evident that the scores after the first two seconds are almost constant. This means that the users in this case are unable to benefit from having a longer response time. It seems that the majority of the students during the longer response time are unable to make proper use of the tools (mainly truth table analysis) that have been introduced during the course.

Figure 8: The Proplog score relative to the response time.

In this respect the difference between the Syllog and Proplog cases are remarkable.

7. Analysis of students' results at their exam

It might be suspected that learning logic simply depends on some specific ability to handle abstract notions, symbolic representation etc. If so, there should be a strong correlation between the ability to handle any two branches of logic. For this reason, it might be interesting to compare the Syllog and Proplog marks at the exam, when the students were evaluated with a mark on a 5-point scale, with 5 as the best mark. The results are shown below.

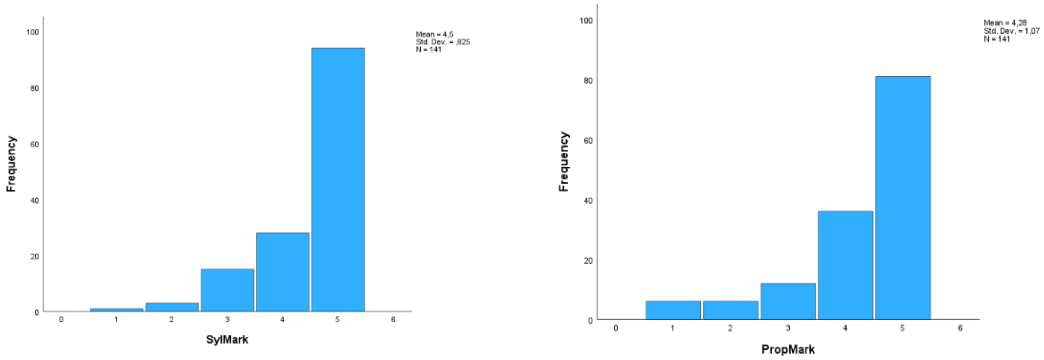


Figure 8: Syllog and Prolog marks at the exam.

Nonparametric Correlation coefficient (Spearman's rho) is 0.54 (p-value < 0.001) between the Syllog and Proplog results. This is a moderate correlation [10, p. 746]. This result certainly does not exclude that there is a specific ability to handle logical problems. On the other hand, logic learning may still depend on some combination of general logic ability, the quality of the courses and the energy invested in course work.

8. Conclusion and perspectives

It is obvious from the above observations that teachers and others can extract a lot of useful information from a careful analysis of the stored log data. The following may be suggested based of the analysis of the log data from the use of Syllog and Proplog during the course:

- The students see it as more difficult to meet the requirements regarding propositional logic than the requirements regarding syllogistics.
- In both cases the invalid arguments are even harder to detect compared with valid ones.
- In average the students can benefit from having more response time when they work with syllogistic arguments. This indicates that the majority of the students in this case are able to use the available tools constructively.
- In average the students can't benefit from having more response time when they work with propositional arguments. This indicates that majority of the students in this case are unable to use the available tools constructively.

- The poor result in propositional logic appears to be related to many students' wrongful identification of the conditional ("if") and the biconditional ("if and only if").

The present study indicates that the teacher during the course should have invested more energy in showing the students have to use to tools designed to evaluate propositional arguments. In addition, the teacher should have focused more on the difference between the conditional and the biconditional. In general, it may be concluded that such observations of response times and other log data may guide the teacher in locating the topics that ought to be more emphasized or adjusted when the course is offered again.

References

1. Long, P.D., Siemens, G.: Penetrating the Fog: Analytics in learning and education. *Edacause Review*, 46(5), 31-40 (2011)
2. Oranje, A., Gorin, J., Jia, Y., & Kerr, D.: Collecting, Analyzing and Interpreting Response Time: Eye-Tracking and Log Data. In: Ercikan, K., Pellegrino, J.W. (eds.) *Validation of Score Meaning For the Next Generation of Assessments: The use of response processes*, pp. 39-51. Routledge: New York (2019)
3. Øhrstrøm, P., Sandborg-Petersen, U., Thorvaldsen, S., Ploug, T.: Teaching Logic through Web-Based and Gamified Quizzing of Formal Arguments. In: Hernández-Leo, D., Ley, T., Klamma, R., Harrer, A. (eds) Scaling up Learning for Sustained Impact. EC-TEL 2013. Lecture Notes in Computer Science, vol 8095, Springer, Berlin, Heidelberg (2013) https://doi.org/10.1007/978-3-642-40814-4_32
4. Øhrstrøm, P., Sandborg-Petersen, U., Thorvaldsen, S., Ploug, T.: Teaching syllogistics through gamification and interactive proofs. In: Conole, G., Klobučar, T., Rensing, C., Konert, J., Lavoué, E. (eds.) Design for Teaching and Learning in a Networked World: Proceedings of EC-TEL 2015. Lecture Notes in Computer Science (LNCS), vol. 9307, Springer (2015). https://doi.org/10.1007/978-3-319-24258-3_70
5. Øhrstrøm, P., Thorvaldsen, S., Sandborg-Petersen, U., Ploug, T.: Teaching Propositional and Syllogistic Logic Using E-learning Tools. In: Rønningsbakk, L., Wu, TT., Sandnes, F., Huang, YM. (eds) Innovative Technologies and Learning. ICITL 2019. Lecture Notes in Computer Science, vol 11937, Springer, Cham. (2019) https://doi.org/10.1007/978-3-030-35343-8_89
6. Øhrstrøm, P., Thorvaldsen, S. & Jakobsen, D.: Learning Analytics Based on Streamed Log Data from a Course in Logic, In: Huang, Y-M. & Rocha, T. (eds.). Innovative Technologies and Learning, Proceedings: ICITL 2023, Lecture Notes in Computer Science, Vol. 14099. pp. 430-440 Springer (2023)
7. Øhrstrøm, P., Ploug, T., Jakobsen, D.: Logisk set. 2nd ed. Metaphysica, Aalborg, (2019)
8. Parry, W., Hacker, E.: Aristotelian Logic. State University of New York Press, (1991)
9. Cohen, L., Manion, L., Morrison, K.: Research Methods in Education (8th ed.). Routledge, (2018).
10. Horn, Laurence R.: From if to iff: Conditional perfection as pragmatic strengthening, Journal of Pragmatics 32, 289-326 (2000)