Teaching Basic Ideas in Logic Using E-learning Tools

Peter Øhrstrøm*1, Ulrik Sandborg-Petersen1, Steinar Thorvaldsen2 and Thomas Ploug1

1 Department of Communication and Psychology, Aalborg University, Rendsburggade14, 9000 Aalborg, Denmark
2 Department of Education, University of Tromsø, 9037 Tromsø, Norway

Abstract

This paper is a study of various e-learning strategies for teaching basic ideas in logic. The focus is mainly on syllogistic validity and deduction. It is a continuation of earlier studies involving practical experiments with students of Communication using the Syllog system, which makes it possible to develop e-learning tools and to do learning analytics based on log-data. The aim of the present paper is to investigate whether the Syllog e-learning tools can be helpful in logic teaching in order to obtain a better understanding of logic and argumentation in general and syllogisms in particular. Four versions of a course in basic logic involving different teaching methods will be compared. Both short- and longterm effects are discussed.

Keywords: Syllogistics, argumentation, learning analytics, logical proofs, deduction, gamified quizzing, logic teaching.

Received on 15 January 2017, accepted on 13 April 2017, published on 20 June 2017

1. Introduction

In this paper we discuss problems and teaching challenges related to courses in basic logic and argumentation offered to 2nd year students in “Communication and Digital Media” at Aalborg University in Aalborg and Copenhagen. The present study is a continuation of previous studies and practical experiments cf. [7], [8], [9], and [10]. Data from the courses in 2012, 2013, 2014 and 2015 will be discussed. The general structure of these courses can be outlined in the following manner:

Period 1. Lectures + homework: General introduction to logic and argumentation. Introduction to classical syllogistics.
Period 2. Lectures + homework: Classical syllogistics (Euler diagrams, Venn diagrams, proofs) and propositional logic (truth tables, proofs).
Period 4. Lectures + homework: Ideas of formal reasoning. The role of logic in everyday life and in scientific argumentation.

About 20 lessons are offered in the course. In addition, the students have to do homework. The total number of students has been 150-200 each year.

During 2012-2015 courses with four different versions of Period 3 have been tested. The tests have focussed on syllogistic reasoning. The aim of the teaching during this period has been to introduce the notions of validity and deductive proofs illustrated in terms of Aristotelian syllogisms.

In order to test and measure the students’ ability to do syllogistic reasoning, the program Syllog has been developed. Sylllog is implemented as a Java-Applet running in the student’s browser, developed using PROLOG+CG (see [3], [4], [5], [6], [13], [14], and [15]). However, the system is not only useful for measuring the students’ ability to do syllogistic reasoning. Versions of the system can also be used in order to support the students in their process of learning the principles of logic.

In Section 2 we present the theory of Aristotelian syllogistics as a deductive system in the classical manner, and it is also explained how the deductive system can be presented in terms of controlled natural language. In Section 3 we present the use of the Syllog system in the
study of the educational problems related to the courses in basic logic. We also discuss two Syllog tools which have been used during the courses. Finally, in Section 4 we compare the outcome of the studies of the four versions of the course and offer suggestions for further research.

### 2. Aristotelian Syllogisms as Deductive Structures

Aristotelian syllogistics has been an essential part of almost all courses in basic logic since the rise of the European university in the 11\textsuperscript{th} century; cf. [11] and [12].

In modern logic teaching the classical (medieval) syllogistics is often presented as a fragment of first order predicate calculus. A classical syllogism corresponds to an implication of the following kind:

\[(p \land q) \supset r\]

where each of the propositions \(p, q,\) and \(r\) matches one of the following four forms:

- \(a(U, V)\) (read: “All \(U\) are \(V\)”)
- \(e(U, V)\) (read: “No \(U\) are \(V\)”)
- \(i(U, V)\) (read: “Some \(U\) are \(V\)”)
- \(o(U, V)\) (read: “Some \(U\) are not \(V\)”)

We may express these functors in terms of first order predicate calculus in the following way:

\[
a(U, V) \iff \forall x: (U(x) \supset V(x)) \\
e(U, V) \iff \forall x: (U(x) \supset \neg V(x)) \\
i(U, V) \iff \exists x: (U(x) \land V(x)) \\
o(U, V) \iff \exists x: (U(x) \land \neg V(x))
\]

The four basic propositions can be related in terms of negation:

\[
i(U, V) \iff \neg e(U, V) \\
o(U, V) \iff \neg a(U, V)
\]

The classical syllogisms occur in four different figures:

- \((u(M, P) \land v(S, M)) \supset w(S, P)\) (1\textsuperscript{st} figure)
- \((u(P, M) \land v(S, M)) \supset w(S, P)\) (2\textsuperscript{nd} figure)
- \((u(M, P) \land v(M, S)) \supset w(S, P)\) (3\textsuperscript{rd} figure)
- \((u(P, M) \land v(M, S)) \supset w(S, P)\) (4\textsuperscript{th} figure)

where \(u, v, w \in \{a, e, i, o\}\) and where \(M, S, P\) are variables corresponding to “the middle term”, “the subject” and “the predicate” (of the conclusion).

In this way, 256 different syllogisms can be constructed. According to classical (Aristotelian) syllogistics, however, only 24 of them are valid. The medieval logicians named the valid syllogisms according to the vowels, \(a, e, i, o,\) involved. In this way the following artificial names were constructed (see [1]):

- 1\textsuperscript{st} figure: barbara, celarent, darii, ferio, barbarix, feraxo
- 2\textsuperscript{nd} figure: cesare, camestres, festino, baroco, camestropon, cesarox
- 3\textsuperscript{rd} figure: darapti, disamis, datisi, felapton, bocardo, ferison
- 4\textsuperscript{th} figure: bramantip, camenes, dimaris, fesapo, fresison, camenop

In these names some of the consonants signify the logical relations between the valid syllogisms, and they also indicate which rules of inference should be used in order to obtain the syllogism in question from the four syllogisms which were considered to be fundamental (i.e. axiomatic): barbara, celarent, darii, ferio (see [1], [5] and [9]).

An even more convincing representation of the deductive system of syllogisms than the one suggested in medieval logic, may be obtained using five fundamental deduction rules. These rules can be formulated symbolically in terms of the conceptual graph interchange format (CGIF) as it was suggested in [9]. However, the rules may also be formulated in terms of a controlled fragment of natural language:

\[
\begin{align*}
\text{(TRANS)} & \quad \text{All } Y \text{ are } Z \\
& \quad \text{Therefore: All } X \text{ are } Z \\
\text{(SUBST)} & \quad \text{All } Y \text{ are } Z \\
& \quad \text{Some } X \text{ are } Y \\
& \quad \text{Therefore: Some } X \text{ are } Z \\
\text{(CONTRA)} & \quad \text{All } X \text{ are } Y \\
& \quad \text{Therefore: All non-} Y \text{ are non-} X \\
\text{(MUT)} & \quad \text{Some } X \text{ are } Y \\
& \quad \text{Therefore: Some } Y \text{ are } X \\
\text{(EX)} & \quad \text{All } X \text{ are } Y \\
& \quad \text{Therefore: Some } X \text{ are } Y
\end{align*}
\]

Note that we allow for negations of terms. The term non-\(X\) is defined as representing all the elements in the universe that are not instants of \(X\). This means that “not \(X\)” is identified with “non-\(X\)” and that “non-non-\(X\)” would be identified with \(X\) (the so-called rule of “double negation”). This means that we can reduce e- and o-propositions in the following way:

\[
\text{“No } X \text{ are } Y \text{”} = \text{“All } X \text{ are non-} Y \text{”} \quad \text{(def. e)} \\
\text{“Some } X \text{ are not } Y \text{”} = \text{“Some } X \text{ are non-} Y \text{”} \quad \text{(def. o)}
\]

The introduction of these definitions entails that, in terms of the controlled natural language, the number of types of propositions in syllogistic reasoning can be reduced from four to two, namely the universal propositions (i.e. “All ... are ...”), and the particular propositions (i.e. “Some ... are ...”). In combination with the option of term negation and the above inference rules we have everything that we
need in order to evaluate all possible syllogisms in classical syllogistics.

It should be noted that (TRANS), which is in fact short for ‘transitivity’, may be read as a version of the syllogism barbara in figure 1, i.e.

All Y are Z
All X are Y
Therefore: All X are Z

(TRANS) may be illustrated graphically in the following manner:

Fig. 1. Illustration of the rule (TRANS).

In terms of sets, Fig. 1 illustrates the idea that if all elements of X belong to Y, and all elements of Y belong to Z, then all elements of X belong to Z. Furthermore, by substituting Z by non-Z we get the syllogism celarent in figure 1 in the following manner:

All Y are non-Z
All X are Y
Therefore: All X are non-Z

The deduction rule (SUBST), which is short for ‘substitution’, can be illustrated in the following manner:

Fig. 2. Illustration of the rule (SUBST).

In terms of sets, Fig. 2 illustrates the idea that if some elements of X belong to Y, and all elements of Y belong to Z, then some element of X belongs to Z. This means that in this case we may substitute Y with Z in “some X are Y”.

It is obvious that (SUBST) in this way leads directly to the syllogism darii in figure 1:

All Y are Z
Some X are Y
Therefore: Some X are Z

If Z is replaced by non-Z we get ferio in figure 1:

All Y are non-Z (= “No Y are Z”)
Some X are Y
Therefore: Some X are non-Z (= “Some X are not Z”)

The three remaining rules are different from the first two in the sense that they only depend on one premise each. (CONTRA) makes it possible to transform a universally quantified proposition, whereas (MUT) makes it possible to transform an existentially quantified proposition. (EX) makes it possible to derive an existentially quantified proposition from a universally quantified proposition.

The deduction rule (CONTRA) may be illustrated graphically using the following diagram:

Fig. 3. Illustration of the rule (CONTRA) – and the rule (EX).

In terms of sets, Fig. 3 illustrates the idea that if all elements of X belong to Y and it is assumed that X is non-empty, then some element of X belongs to Y.

The definition of e mentioned above means (in combination with the rule of double negation) that the deduction rule (CONTRA) can work in four different ways:

All non-X are non-Y
Therefore: All Y are X

All non-X are Y
Therefore: All non-Y are X

All X are non-Y
Therefore: All Y are non-X

All X are Y
Therefore: All non-Y are non-X

Fig. 3 may in fact also illustrate the rule (EX), in the sense that if all elements of X belong to Y and it is assumed that X is non-empty, then some element of X belongs to Y.
The deduction rule (MUT) may be illustrated graphically using the following diagram:

![Fig. 4. Illustration of the deduction rule (MUT).](image)

Fig. 4 clearly shows that if some X belongs to Y then some Y belongs to X.

The above inference rules may be understood as the basis of a deductive system that makes it possible to derive the conclusion in a syllogistic argument from its premises if and only if the syllogistic argument under consideration is a valid syllogism. It is easy to present (TRANS), (SUBST), (CONTRA), (MUT) and (EX) in terms of Euler circles. In this way it can be made clear that the rules are intuitively reasonable.

Using this deductive approach to the syllogisms, it is possible to show a number of interesting results concerning the invalidity of certain syllogistic arguments. For instance, by going through the five rules of inference it is evident that if both premises are existential, then nothing new follows regarding the relation between subject and predicate. The same holds if both premises are negative, i.e., \(\neg\)-propositions or \(\neg\)-propositions.

The use of the inference rule (EX) has sometimes been seen as controversial, and the 9 syllogisms which depend on this rule have consequently been seen as “questioned”. As indicated above (EX) only works on the condition that the sets in question are non-empty. Clearly, the rule has to be rejected, if we hold that the statement “all S are P” is true given that S is the empty set. Therefore, if this is accepted it should obviously not be permitted to deduce “some” from “all”. If the EX rule is excluded, the number of valid syllogisms is reduced from 24 to 15.

3. Teaching Syllogistics Using Syllog

The Syllog system generates syllogisms at random, and the user is supposed to evaluate them using the system. The activities of the students when working with the system are logged, and the log-data from the use of the system may give rise to very interesting learning analytics. Fig. 5 shows the interface of one the Syllog versions:

![Fig. 5. Gamified quizzing with Syllog.](image)

A student’s ability to do syllogistic reasoning can be analysed in terms of the score calculated on the basis of log-data from the use of Syllog. This score is calculated as:

\[
\text{Score} = \frac{\text{correct answers}}{\text{answer count}}
\]

The score measures how well the user is doing in evaluating the validity of syllogistic arguments. Teaching of logic is at least in part aimed at raising this score. The statistical analyses of the scoring data were performed using standard methods from descriptive statistics and statistical testing. An interesting question concerns the students’ ability to evaluate the validity of syllogisms before receiving formal training on this subject [7, 8, 9]. In the previous studies we have provided evidence to the effect that the students’ ability to distinguish between valid and invalid syllogisms before the teaching starts is significantly higher than the level of guessing. The value of this early score appears to be rather stable from year to year during the period 2012-15. The studies suggest 0.608 as the value of this early score.

Using our data we have been able to find the invalid syllogisms that obtain the lowest scores i.e. the most remarkable errors made by the students. The result regarding the invalid syllogisms is that the following three syllogisms obtain the lowest score:

- Fig. 1: iio
- Fig. 3: aee
- Fig. 4: iao

The result regarding the valid syllogisms is:

- Fig. 4: aeo (camenop)
- Fig. 2: aeo (camestrop)

† This estimate was in fact measured with a similar group of 2nd year Communication students at Aalborg University in 2016. In this case it was found that the students produced 2260 correct answers out of 3786, corresponding to an average score of 0.597. This gives further support to the claim that the average score for newcomers from Communication students is very close to 0.60.
Fig. 1: eao (feraxo)

In this way, these syllogisms appear to be the three most difficult (invalid or valid) syllogisms to evaluate. It is remarkable that all six syllogisms mentioned above have negative conclusions. It should also be noticed that the three syllogisms mentioned above belong to the group of 9 syllogisms that do not depend on the (EX) rule.

It is easy to see why the students could be wrong in their evaluation of the syllogisms mentioned above. Consider, for instance, the following presentation of the iio-syllogism in Fig. 1:

Some parents are males.
Some doctors are parents.
Ergo:
Some doctors are not males.

Obviously, all statements involved in this syllogism are true relative to the real world. In order to realise the invalidity of this syllogism, the students have to imagine a counterfactual state-of-affairs according to which the premises are true, but the conclusion false (i.e. a possible world in which all doctors are males). At least for some students this appears to be a rather difficult task.

This kind of information is clearly valuable for teachers who want to design a course in basic logic. However, it is certainly also interesting to measure the average score after some logic teaching. The study based on data from the 2012 version of the course showed that there is no or very limited improvement in the score if it is measured after a traditional course in basic logic (with traditional work with exercises on paper during Period 3).

No significant improvement of the average score was detected (see [7]).

Fig. 5: The user interface shown in Fig. 5 it should be noted that “The number of correct answers in a row” is displayed. Using this facility it is possible to establish a competition between the groups of students, and this rather simple gamification element actually turns out to work as a motivation in the practical setting. This effect of simple gamified quizzing was studied based on the data from the course in 2013 and further studies in 2015. This study showed that the use of gamification elements can have some positive effects on the motivation to learn, and in combination with a traditional course on syllogistics it can lead to an increased understanding of logical validity in the sense that the student’s ability to evaluate the validity of an arbitrary syllogism becomes better (see [8] and [10]).

During Period 3 of the 2014 version of the course, the students could do exercises in small groups using a version of Syllog including the deduction rules presented in section 2. The rules were presented on the screen in terms of the CGIF formalism after a general introduction to the formalism (a lecture). The gamification facility mentioned above was also included in the interface. The study showed that only a small fraction of students could benefit from the use of this system. No significant improvement of the average score was detected (see [9]).

During Period 3 of the 2015 course, the students could work in small groups with a proof facility based on the five deduction rules mentioned in Section 2, presented in terms of controlled natural language. The user interface is shown in Fig. 6. The user can click on New to get a new syllogism presented on the screen. Then the user may apply some of the inference rules (‘Trans’, ‘Subst’, ‘Contra’, ‘Mut’, ‘Ex’) to see what follows from the two premises and from other propositions that have been proved so far. This is done by clicking on the button corresponding to the inference rule that the student wants to apply. Whenever ready the user may decide whether he or she believes the syllogism to be valid or invalid. This is done by clicking on the relevant button. In this way the user may perform experiments with the syllogisms in question. Hopefully, this leads to a deeper understanding of syllogistic validity.

Fig. 6. The interface of the Syllog system used in 2015.

As indicated in the above Fig. 6, the system automatically translated the premises of the argument into a controlled fragment of natural language. E.g. the premise “No parents are doctors” is immediately translated into “All parents are non-doctors”, etc. In this way it becomes easier for the user to see which of the rules (if any) may be applied.

The proof (or demonstration) produced in Fig. 6 may be rephrased with addition of some explanations in the following manner using a kind of so-called natural deduction:

(1) No parents are doctors (premise)
(2) Some males are doctors (premise)
(3) All parents are non-doctors (1, def. e)
(4) All doctors are non-parents (3, CONTRA)
(5) Some males are non-parents (2, 4, SUBST)
(6) Some males are not parents (5, def. o)
The interesting question from a learning perspective is whether the students have been able to benefit from the introduction of this proof procedure and from the use of the implementation of it in Sylog. As we shall see in the concluding section, this question has been investigated experimentally.

The average score was measured at the beginning Period 3 when the work with the deduction module starts (and after some work with “gamified quizzing”). After Period 3 the average score was measured again. The aggregated results and the scores are shown in Table 1. The results support strong statistical evidence against the presumption that student will handle the syllogisms equally well before and after Period 3 (p-value < 0.001 by the $\chi^2$ test).

Table 1. The 2x2 table summarizing counts from the 2015 course of how often students replied correctly to the syllogisms in the beginning of Period 3 and immediately after this period. These values may be compared with the value (from earlier studies) of the score before the teaching starts, i.e. 0.608.

<table>
<thead>
<tr>
<th>Correct reply?</th>
<th>Yes</th>
<th>No</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>The beginning of Period 3 (n=133)</td>
<td>1145</td>
<td>615</td>
<td>0.651</td>
</tr>
<tr>
<td>After Period 3</td>
<td>1112</td>
<td>462</td>
<td>0.706</td>
</tr>
</tbody>
</table>

On Dec. 1, 2015, more than 8 months later, 12 students participated in a follow-up of the same experiment in order to evaluate to what extent the improved skills are lasting. The data are shown here:

Table 2. The results of the same experiment as in Table 1 carried out 8 months later.

<table>
<thead>
<tr>
<th>Correct reply?</th>
<th>Yes</th>
<th>No</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>After 8 months</td>
<td>195</td>
<td>130</td>
<td>0.628</td>
</tr>
</tbody>
</table>

Compared with the results in Table 1 the results in Table 2 show a significant drop in score from 0.706 to 0.628 (p-value=0.0002 by the $\chi^2$ test). However, the scores did not drop to the pre-course level of 0.608, even if the difference is not significant (p-value=0.81 by the $\chi^2$ test). This long term evaluation is based on only 12 students, but may indicate that our proof system is mainly for the short term memory.

The results indicate that although many students during the course apparently have obtained a somewhat clear understanding of logical validity as a purely abstract and formal notion, this knowledge is partly forgotten after the course. The so called Ebbinghaus Forgetting Curve is often used to demonstrate how we forget new information that we don’t work with repeatedly [2], and the result in Table 2 is consistent with this theory. On the other hand, it should be admitted that more information is needed in order to give a full account of the causes behind the measurements. Among other avenues of research, it would be interesting to know more precisely to what extent the (apparent) effects from the course regarding the understanding of logical validity are lasting. Furthermore, it will also be important to study how we can obtain more lasting effects of the teaching of logic.

4. Conclusions

The present study as well as the previous studies provides strong evidence for the usability of the log-functionality of PROLOG+CG in order to establish relevant analytics regarding the teaching and learning of logic.

Based on the study of the Sylog data from the courses in 2013 and 2014 we have seen that we may benefit from the use of interactive e-learning tools during a logic course, whereas no significant improvement of the ability to do syllogistic reasoning could be detected after the traditional course offered in 2012. Our study provides evidence that students during a course using such a system improve their ability to evaluate logical validity significantly. In particular, the student could benefit from having access to Sylog with deduction rules in terms of natural language, whereas a similar system in terms of CGIF (as in the 2014 course) was of almost no use in most cases. Table 3 comparatively summarises the results from the four versions of the course.

Table 3. A comparison of the results based on log-data from the four versions of a course in basic logic. Only the content of Period 3 of the course has been changed from year to year.

<table>
<thead>
<tr>
<th>Year</th>
<th>Period 3</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>Traditional work with logic exercises (no use of e-learning tools).</td>
<td>No significant improvement of the ability to do syllogistic reasoning was detected after the course [6].</td>
</tr>
<tr>
<td>2013</td>
<td>Traditional work with logic exercises + gamified quizzing with Sylog.</td>
<td>A small but significant effect of the teaching was detected [7].</td>
</tr>
<tr>
<td>2014</td>
<td>Traditional work with logic exercises + gamified quizzing with Sylog + work with a deduction module in terms of CGIF.</td>
<td>Mixed result. Only some of the students could benefit from the work with the CGIF module. No significant improvement of the average ability to do syllogistic reasoning was detected [8].</td>
</tr>
<tr>
<td>2015</td>
<td>Traditional work with logic exercises + gamified quizzing with Sylog + work with a deduction module in terms of natural language.</td>
<td>A significant improvement of the ability to do syllogistic reasoning was detected. The highest value of the average score (0.706) was measured in this case.</td>
</tr>
</tbody>
</table>

Observations of the students during their work with the tools suggest that the work with the tools in 2013 and in
2015 stimulated their motivation and interest in the topic. Fig. 7 is an attempt to put the results in perspective:

![Score Chart]

Fig. 7. Scores obtained in Test 1 (2012, without the use of e-learning tools), Test 2 (2013, repeated in 2015, with a gamified quizzing tool) and Test 3 (2015, with a gamified quizzing tool and a tool for the investigation of the deductive structures of syllogisms).

The 2014 study [9] shows that most of the Communication students were unable to benefit from a Syllog tool using CGIFs in order to investigate the deductive structure of the syllogisms. It is likely that this will be the case for any tool that makes use of a complex formalism or symbolic language. This assumption is supported by the finding of the present study that the students can in fact benefit from the use of the deductive module presented in terms of controlled natural language. This module probably leads to a better understanding of syllogistic validity. Still, an average score of 0.706 in the 2015 test is not very impressive. Obviously, the challenge is to develop better e-learning tools and teaching strategies in order to improve the students’ ability to do syllogistic reasoning even more. For this purpose, it might be useful to know more about the kind of difficulties that the students are facing when they are working with syllogisms. Further studies of the log-data may provide such information, and we may thereby obtain useful information on how to develop more effective e-learning tools and teaching strategies.

There are a number of other further studies that would be interesting to carry out. One of them would focus on the question of individualised learning. Firstly, it may be interesting to require the students to use the tool individually rather than letting them do it groups. In this way the volume of data will be increased which may allow more fine grained measurements. In addition, it will be interesting to investigate how the performance of the students is affected by replacing interpersonal dialogue and discussion with inner reflection. Secondly, there may be patterns in the individual student’s ability and inability to evaluate logical validity successfully. Certain syllogistic forms may be more difficult for the individual student to evaluate as to their validity. Interactive learning modules may be used to identify such patterns, and subsequently be targeted at improving the individual patterns in the inability to evaluate logical validity.

Another interesting further research project would be the study of motivational effects. The use of interactive learning modules for logic teaching may have both short- and long-term effects on students’ motivation to learn about logics. We have mainly investigated the short-term effect of using learning modules for the students’ ability to evaluate logical validity. Although they have limited effects on this particular ability, they may have an effect on students’ motivation to learn about logics, and this may ultimately have an effect on the students’ long-term development of logical reasoning skills.

References


