Teaching Syllogistics Using Conceptual Graphs

Peter Øhrstrøm[1], Ulrik Sandborg-Petersen[1],
Steinar Thorvaldsen[2], and Thomas Ploug[1]

[1] Department of Communication and Psychology, Aalborg University,
9000 Aalborg, Denmark
{poe,ulrikp,ploug}@hum.aau.dk

[2] Department of Education, University of Tromsø, 9037 Tromsø, Norway
steinar.thorvaldsen@uit.no

Abstract. It has for centuries been commonly believed that syllogistic reasoning is an essential part of human rationality. For this reason, Aristotelian syllogistics has since the rise of the European university been a standard component of logic teaching. During the medieval period syllogistic validity was presented in terms of a number of artificial words designed to summarize the deductive structure of this basic system. The present paper is a continuation of earlier studies involving practical experiments with informatics students using a student-facing Java-Applet running in the student's browser, implemented using the Prolog programming language as embodied in a Java implementation called Prolog+CG. The aim of the present paper is to study some interesting conceptual aspects of syllogistic reasoning and to investigate whether CG formalism can be helpful in order to obtain a better understanding of syllogistic reasoning in general and the system of Aristotelian syllogisms conceived as a deductive (axiomatic) structure in particular. Some prototypes of tools for basic logic teaching have been developed using Prolog+CG, and various preliminary tests of the tools have been carried out.

Keywords: Syllogistics, argumentation, conceptual graphs, deduction, logic teaching.

1 Introduction

Aristotelian syllogistic has been an essential part of almost all courses in basic logic since the rise of the European university in the 11th century; cf. [1] and [4]. This kind of reasoning has commonly been believed to constitute a crucial component of the ideal of basic human argumentation as it is carried out in social life. Clearly, most of us make frequent mistakes in the way we are treating syllogistic forms in our discussions in daily life. However, normally we are able to recognize and understand what went wrong when we are confronted with our errors. For this reason, we very much treasure and appreciate the logical ideal of correctness which has been expressed in syllogistic theory as it has been known since Aristotle.

It is an important aspect of Aristotelian syllogisms that the valid syllogisms can be reduced or explained in terms of some even more basic principles. In fact, the valid

N. Hernandez et al. (Eds.): ICCS 2014, LNAI 8577, pp. 217–230, 2014.
DOI: 10.1007/978-3-319-08389-6_18,

syllogisms may be understood as an axiomatic system. Actually, the system of Aristotelian syllogisms may be understood as the very first presentation of an axiomatic system in logical literature. This view of the syllogisms as a deductive system was in fact incorporated in the standard medieval presentation of the syllogisms. In modern logic there is also a strong emphasis on the importance of deduction. John Sowa has suggested that deductive reasoning is represented in terms of so-called conceptual graphs cf. [5]. This formalism may be presented using the so-called Conceptual Graph Interchange Format (CGIF); cf. [7]. In the present paper we explore the use of this option in logic teaching. The main questions are:

a) How should the understanding of the syllogisms as a deductive system be incorporated in modern logic teaching?
b) Can a better understanding of the syllogisms as deductive system give rise to a more correct syllogistic reasoning in practice?

The present study is a continuation of earlier studies and practical experiments involving informatics students and other students from the humanities; cf. [9] and [10]. The experiments have been carried out using the so-called Syllog system, i.e. a Java-Applet running in the student's browser, implemented using the Prolog programming language as embodied in a Java implementation called Prolog+CG; cf. [2], [3], [6] and [8]. Syllog generates syllogisms at random, and the user is supposed to evaluate them using the system.

In section 2 of this paper we present the theory of Aristotelian syllogistics as a deductive system in the classical manner as well as a more modern way in terms of conceptual graphs. In section 3 we describe the general experimental setup as well as the practical experiments which have been carried out. In section 4 we discuss the results of the experiments. Section 5 is based on a qualitative interview with a group of students, and it deals with a potential improvement of the experimental setup which can lead to the development of a tool which may be used in logic teaching in order to give the students a better understanding of the nature of deductive structures. In section 6 it will be discussed what we may conclude from the present study and also what further developments and investigations the present study may suggest.

2 Aristotelian Syllogisms as Deductive Structures

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

$$(p \wedge q) \supset r$$

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”)

These four functors were suggested by the medieval logicians referring to the vowels in the words “affirmo” (latin for “I confirm”) and “nego” (latin for “I deny”), respectively. Given that U and V are interpreted as predicates, we may express the functors in terms of the usual first order predicate calculus in the following way:

$$a(U, V) \leftrightarrow \forall x: (U(x) \supset V(x))$$
$$e(U, V) \leftrightarrow \forall x: (U(x) \supset \sim V(x))$$
$$i(U, V) \leftrightarrow \exists x: (U(x) \wedge V(x))$$
$$o(U, V) \leftrightarrow \exists x: (U(x) \wedge \sim V(x))$$

These four basic propositions are related in terms of the negation in the following manner:

$$i(U, V) \leftrightarrow \sim e(U, V)$$
$$o(U, V) \leftrightarrow \sim a(U, V)$$

The classical syllogisms occur in four different figures:

$$(u(M, P) \wedge v(S, M)) \supset w(S, P) \quad \text{(1st figure)}$$
$$(u(P, M) \wedge v(S, M)) \supset w(S, P) \quad \text{(2nd figure)}$$
$$(u(M, P) \wedge v(M, S)) \supset w(S, P) \quad \text{(3rd figure)}$$
$$(u(P, M) \wedge v(M, S)) \supset w(S, P) \quad \text{(4th figure)}$$

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). The premises are mentioned according to the convention that the grammatical subject (here S) of the conclusion should appear in the second of the premises.

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]):

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

If needed, we may add the figure number to the name (e.g. baroco2, fesapo4). 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 syllogisms which were considered to be fundamental: barbara, celarent, darii, ferio. – In fact, the system of syllogisms may in this way be seen as an axiomatic system with these four axioms and the following inference rules (see [1] and [3]):

x-rule-1:	$a(U,V) \rightarrow i(U,V)$	$\forall x: (U(x) \supset V(x)) \rightarrow \exists x: (U(x) \wedge V(x))$
x-rule-2:	$e(U,V) \rightarrow o(U,V)$	$\forall x: (U(x) \supset \sim V(x)) \rightarrow \exists x: (U(x) \wedge \sim V(x))$
s-rule-1:	$i(U,V) \leftrightarrow i(V,U)$	$\exists x: (U(x) \wedge V(x)) \leftrightarrow \exists x: (V(x) \wedge U(x))$
s-rule-2:	$e(U,V) \leftrightarrow e(V,U)$	$\forall x: (U(x) \supset \sim V(x)) \leftrightarrow \forall x: (V(x) \supset \sim U(x))$
m-rule:	$(p \wedge q) \supset r \leftrightarrow (q \wedge p) \supset r$	
c-rule-1:	$(p \wedge q) \supset r \leftrightarrow (\sim r \wedge q) \supset \sim p$	
c-rule-2:	$(p \wedge q) \supset r \leftrightarrow (p \wedge \sim r) \supset \sim q$	

In addition, p-rule-1 may be defined as a combination of x-rule-1 and s-rule-1, whereas p-rule-2 may be defined as a combination of x-rule-2 and s-rule-2. It should also be noted that x-rule-1/2 is used from left to right in the conclusion, whereas it is used from right to left in the premises.

The use of the axioms in the deduction is rather simple, since only one axiom is involved in each deduction. From barbara1 we may derive barbari**x**1 (by x-rule-1), baro**c**o2 (by c-rule-2), bo**c**ardo3 (by c-rule-1), bra**m**anti**p**4 (by m-rule and p-rule-1).

From celarent1 we may derive ce**s**are2 (by s-rule-2), ca**m**e**s**tre**s**2 (by m-rule and two applications of s-rule-2), ca**m**e**s**tro**p**2 (by m-rule, s-rule-2 and p-rule-2), ce**s**aro**x**2 (by s-rule-2 and x-rule-2), ca**m**ene**s**4 (by m-rule and s-rule-2), ca**m**eno**p**4 (by m-rule and p-rule-2). From darii1 we may derive dara**p**ti3 (by p-rule-1), di**s**a**m**i**s**3 (by m-rule and two applications of s-rule-1), dati**s**i3 (by s-rule-1), di**m**ari**s**4 (by m-rule and s-rule-1). From ferio1 we may derive fera**x**o1 (by x-rule-1), fe**s**tino2 (by s-rule-2), fela**p**ton3 (by p-rule-1), feri**s**on3 (by s-rule-1), fe**s**a**p**o4 (by s-rule-2 and p-rule-1), fre**s**i**s**on4 (by s-rule-2 and s-rule-1).

There is obviously a great deal of structural beauty built into this system of artificial names. It should be noted that the names of a syllogism may in fact be read as an abbreviation of its proof. For instance, the proof of ca**m**e**s**tro**p**2 from celarent1 may be presented in the following manner:

$(e(M, P) \wedge a(S, M)) \supset e(S, P)$	(celarent1)
$(a(S, M) \wedge e(M, P)) \supset e(S, P)$	(use of m-rule)
$(a(S, M) \wedge e(P,M)) \supset e(S, P)$	(use of s-rule-2)
$(a(S, M) \wedge e(P,M)) \supset o(P,S)$	(use of p-rule-2; ca**m**e**s**tro**p**2; QED)

Note that in this demonstration of ca**m**e**s**tro**p**2, *P* plays the role of the grammatical subject in the conclusion. For this reason the order of the premises should according to the medieval standard be presented as it is done here.

In logic teaching this nice medieval system may obviously be used in order to illustrate the idea of a logical proof. In this way, the system of the 24 valid syllogisms may be reduced to a system of four axioms and seven rules of inference. From a modern perspective we have to say that this reduction does not go far enough. On the other hand, it should be admitted that the system is elegant, and that the idea of using the proof of a theorem as its name is rather remarkable.

Given that we want to produce an even more convincing reduction of the system of syllogisms than the one which was suggested in medieval logic, we may apply conceptual graphs represented in the conceptual graph interchange format (CGIF); cf.

[7]. In terms of this formalism the four basic propositions syllogistics may be represented in the following manner, where "–attr->" stands for "has the attribute of":

a(U, V)↔[All: *x] [If: (?x –attr-> U) [Then: (?x-attr-> V)]]
e(U, V) ↔[All: *x] [If: (?x –attr-> U) [Then: ~(?x-attr-> V)]]
i(U, V) ↔ [*x] (?x –attr-> U) (?x-attr-> V)
o(U, V) ↔[*x] (?x –attr-> U) ~(?x-attr-> V)

In the following, we take the liberty of ignoring '*' and '?' in the CGIFs. - The CG formalism is to some extent integrated in the PROLOG+CG. In fact, expressions like, [x] –attr-> [U], are allowed according to the syntax of this system. For instance, if we want to represent "All Greeks are Mortal" in terms of CGs one might consider the following PROLOG clause:

[x] -attr-> [Mortal] :- [x] -attr-> [Greeks].

This representation will, however, be too simplistic in our case. We need more freedom of expression concerning the proof system. It may be argued that the representation of the syllogisms as implications should be represented in terms of sequents, i.e.

$p, q \vdash r$

This representation indicates that *r* is derivable from the combination of *p* and *q*. What derivability means in this context then has to be specified in terms of some rules of inference. The following rules are suggested to cover the deductive system of the Aristotelian syllogisms:

(TR) [All: x] [If: G_1(x) [Then: G_2(x)]]
[All: x] [If: G_2(x) [Then: G_3(x)]]
Therefore:
[All: x] [If: G_1(x) [Then: G_3(x)]]

(SUB) [All: x] [If: G_1(x) [Then: G_2(x)]]
[x] G_1(x) G_3(x)
Therefore:
[x] G_2(x) G_3(x)

(CON) [All: x] [If: G_1(x) [Then: G_2(x)]]
Therefore:
[All: x] [If: ~G_2(x) [Then: ~G_1(x)]]

(MUT) [x] G_1(x) G_2(x)
Therefore:
[x] G_2(x) G_1(x)

(EX) [All: x] [If: G_1(x) [Then: G_2(x)]]
Therefore:
[x] G_1(x) G_2(x)

The usual understanding of negation and double negation is also assumed i.e. for any graph, G, it holds that: ~~G ↔ G. Given these rules any of the 24 valid syllogisms may be demonstrated. Consider, for instance, the syllogism camestrop2 mentioned above. In this case we have to prove:

[x] (x –attr-> P) ~(x-attr-> S)

Given

[All: x] [If: (x –attr-> S) [Then: (x-attr-> M)]]
[All: x] [If: (x –attr-> P) [Then: ~(x-attr-> M)]]

The proof can be carried out in the following manner:

1. [All: x] [If: (x –attr-> S) [Then: (x-attr-> M)]] (premise)
2. [All: x] [If: (x –attr-> P) [Then: ~(x-attr-> M)]] (premise)
3. [All: x] [If: (x –attr-> M) [Then: ~(x-attr-> P)]] (from 2 & CON)
4. [All: x] [If: (x –attr-> S) [Then: ~(x-attr-> P)]] (from 1, 3 & TR)
5. [All: x] [If: (x –attr-> P) [Then: ~(x-attr-> S)]] (from 4 & CON)
6. [x] (x –attr-> P) ~(x-attr-> S) (from 5 & EX; QED)

Note that here not only camestrop2 has been demonstrated. 1-5 actually also constitute a proof of camestres2.

The above demonstration is in fact a relatively complicated proof. In many of the other cases, the numbers of steps in the proofs are much smaller. Consider for instance darii1:

1. [All: x] [If: (x –attr-> M) [Then: (x-attr-> P)]] (premise)
2. [x] (x –attr-> S) (x-attr-> M) (premise)
3. [x] (x –attr-> S) (x-attr-> P) (from 1, 2 & SUB; QED)

Using this deductive approach to the syllogisms, it will also be possible to show important results concerning the invalidity of certain syllogistic arguments. For instance, by going through the rules of inference listed above it is evident that if both premises are existential, then nothing follows. The same holds if both premises are negative i.e. o-propositions or e-propositions.

The use of (EX) corresponding to the medieval x-rule has sometimes been seen as controversial and the 9 syllogisms which depend on this rule have consequently been seen as “questioned”. Clearly (EX) has to be rejected, if we accepted that the statement “all S are P” can be true even if the concept S corresponds to the empty set.

3 The Experiments in the Present Study

From the earlier studies, [9] and [10], it can be concluded that we have very strong evidence for the following claims regarding the way syllogisms are handled by students before they have been introduced to the theories of syllogisms:

a) In many cases students can make correct distinctions between valid and invalid syllogisms. Their ability to do so is significantly higher than the level of guessing.
b) It is easier for the students to find the correct validity in case of a valid syllogism than in case of an invalid syllogism, i.e. there are more errors in spotting the invalidity of invalid syllogistic arguments than in spotting the validity of valid syllogisms.
c) If the conclusion of the syllogistic argument is true (i.e. corresponds with the real world) then the students are likely to assume that the argument is valid.
d) The students are more likely to accept the validity of the 15 "unquestioned" syllogisms than the validity of the 9 "questioned" syllogisms.
e) In [10] it has been studied how gamified quizzing can be used in logic teaching. This study shows that at least in some cases the responses from the students indicate that the use of gamification had some positive effects on the motivation to learn.

The earlier studies also deal with the question of possible improvements in the score caused by the logic teaching. Here the main conclusion can be stated in the following manner:

f) There is no or just a very small improvement in the score if it is measured after the presentation of the medieval theory of syllogisms in general.

In the earlier studies we did not include all 232 invalid syllogistic arguments, but we selected 24 invalid syllogistic arguments assumed to be somewhat "tempting". This approach has been challenged. In consequence we have in the present study included all 232 invalid syllogisms. The present version of the system, Syllog, first selects at random whether the next syllogism to present on the screen should be valid or invalid. If it is supposed to be valid then one of the 24 valid syllogisms is selected at random. If it is supposed to be invalid then one of the 232 invalid syllogisms is selected at random. In each case the user is asked to evaluate the argument presented on the screen.

The 2nd year students of Humanistic Informatics at Aalborg University (in Aalborg and in Copenhagen) were on different dates in February 2014 asked to run Syllog in groups of 2-4 before they were introduced to syllogistic logic. All test results have been logged by Syllog. The score was computed as

Score = correctanswers/answercount

The statistical analyses of the scoring data were performed using standard methods from descriptive statistics and statistical testing. The two sample t-test is applied to detect increased score, and to look for significant differences between results from the pre-test and the post-test. The chi-square test is applied to detect differences in how students handle invalid and valid syllogisms.

In table 1 we compare the results from the pre-test (i.e. before the lectures on syllogisms) in Aalborg using the old dataset from earlier studies, with the test in Aalborg using the new dataset, where all invalid syllogisms are represented. There is no significant difference in how student handle the two datasets, and the average exercise score of the old dataset is 0.564, compared to 0.557 in the new data.

Table 1. The 2x2 tables summarizing counts of how often previous and present students in Aalborg replied correctly to invalid syllogisms in the pre-test. The p-value is by the chi-square test.

	Correct reply? Pre-test	
	No	Yes
Old dataset (24 invalid syllogisms)	**538**	**697**
New dataset (232 invalid syllogisms)	**627**	**789**
p-value	0.74	

This means that in this context it makes no essential difference to include all 232 invalid syllogistic arguments in Syllog. The procedure used in the earlier studies turns out to be fully satisfactory, when it comes to measuring the ability to do valid syllogistic reasoning.

The students were also asked to run the program after the lectures on syllogistic logic. The lectures dealt with the medieval theory of syllogism as well as a brief introduction (about 30 min.) to the treatment of syllogisms in terms of conceptual graphs (CGIF) focusing on the CG formalization of the syllogistic propositions and the rules of inference mentioned above.

In the pre-test the students were just presented with the arbitrary syllogism in its verbal form, e.g.

All teachers are mothers.
All teachers are females.
Ergo: Some females are mothers.

In the post-test, the students were also given the translations of the premises and the conclusion into conceptual graphs (CGIF). With the above example that would mean that the following translated are presented on the screen:

[All: x] [If: (x –attr-> [teachers]) [Then: (x-attr-> [mothers])]]
[All: x] [If: (x –attr-> [teachers]) [Then: (x-attr-> [females])]]
Ergo: [x] [(x –attr-> [females]) (x -attr-> [mothers])]

In addition, the system offered the possible deductions from the premises using (EX). This means that the following derivations were shown on the screen as well:

[x] [(x –attr-> [teachers]) (x -attr-> [mothers])]
[x] [(x –attr-> [teachers]) (x -attr-> [females])]

The challenge for students would then be to determine whether (TR), (SUB), (CON) and (MUT) can be used in order to derive the conclusion from the premises. In this case it should be noted that the conclusion obviously follows by (SUB) from the second premise and the first of the above derivations.

The pre-test were performed by 110 groups in Aalborg and 52 groups in Copenhagen. The post-test were performed by 49 groups in Aalborg, and 36 groups in Copenhagen. This big dropout in Alborg creates a bias for our analyses that must be handled with care. In fact we have decided to ignore the Aalborg Post-test and to concentrate on the results from Copenhagen in this case. All results are, however, shown in Table 2.

Table 2. The above table are summarizing the pre-test versus post-test results of the students average group score in Aalborg and Copenhagen. The p-values are by the two sample t-test (one sided not assuming equal variance). The analyses gives no statistical evidence against the presumption that student will handle syllogism equally well before and after the lecture on syllogisms and conceptual graphs.

		Mean Score	**SD Score**	**p-value Pre vs. Post**
Aalborg	**Pre (N=110)**	0.633	0.191	0.265
	Post (N=49)	0.598	0.176	
Copenhagen	**Pre (N=52)**	0.590	0.259	0.096
	Post (N=36)	0.677	0.240	

Table 2 shows inconsistent results with respect to progress between the pre-test and the post-test. The achieved scores are observed to be between 0.59 and 0.68, and are all significantly higher than random guessing with an expected score of 0.5 (p-values 0.01 or lower by the one sample t-test).

The below figure 1 shows the distribution of the scores in the post-test. It seems that there is a certain tendency of “two or more tops” in the distribution.

We also compared how much time the student groups spent thinking on each task of the syllogism exercise. The results are shown in Table 3, which shows that the students spent significantly more time to think about each task after the lecture had been given.

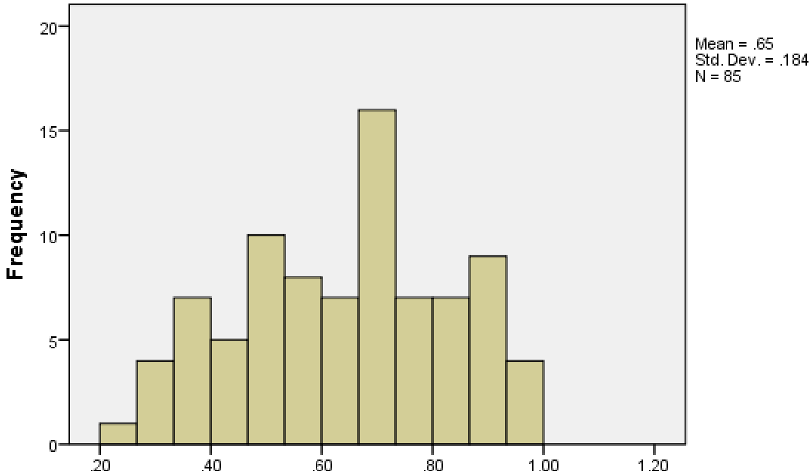

Fig. 1. The distribution of scores in the post-test for the participating groups (N=85)

Table 3. Table are summarizing the pre-test versus post-test results of students time spent on each task of the exercise. The p-values are by the two sample t-test (two sided not assuming equal variance). Some statistical evidence are found against the presumption that student will handle syllogism equally fast before and after the lecture.

		Mean time (s)	**SD time (s)**	**p-value Pre vs. Post**
Aalborg	**Pre (N=110)**	30.9	16.6	0.040
	Post (N=49)	39.2	25.4	
Copenhagen	**Pre (N=52)**	49.8	34.7	0.098
	Post (N=36)	61.3	32.3	

4 Discussion of the Experimental Results

It is evident that the students use more time on each syllogistic argument in the post-test than in the pre-test. This suggests that the presented CGs make them reflect. There is, however, no significant improvement of the average score when the CGs are listed along with the textual presentation of the argument. It should, however, be noted that Figure 1 shows certain tendency towards “two or more tops”. This may indicate that there could be a group of students who are, in fact, able to benefit from the CGs although the majority of students do not benefit from the display of CGs.

The fact that there is no significant improvement of the average score when the CGs are listed as well may not be too surprising given the results from the earlier studies which have been summarised in e) above. In fact, it seems to be very difficult to obtain any improvement in the score through logic teaching. One possibility is that the logical part of the human mind is only trainable within a quite narrow limit, when it comes to practical skills in everyday reasoning. Even if this turns out to be the case, it does not necessarily mean that logic teaching is useless. As one of the students stated during the discussion, based on the insights from the teaching the students may better understand what is wrong in the invalid syllogistic arguments and why the valid arguments hold. In this way, although their reasoning skills in daily life are pretty much the same after the logic lectures, the students may in fact from the logic classes have gained a much deeper understanding of the very notion of validity. On the other hand, it may still be possible to use the CGs in a more convincing manner in order to give the students an even deeper understanding of what valid reasoning is. For this reason, it might still be a good idea to look for designs and tools different from the present post-test through which the students can benefit from the structural qualities of conceptual graphs in an optimal manner. One problem seems to be that the present design with all the translations and derivations mentioned above is too overloaded and complicated, which means that it has become hard to read and difficult to use.

5 Can Conceptual Graphs be Useful in Logic Teaching?

Would it be possible to improve the experimental setup used in the Post-test in order to create a design or even a tool which can effectively support the students in their understanding of syllogistic reasoning? In order to answer this question a qualitative interview with a group of students has been carried out. We conducted a semi-structured focus group interview with 7 students following the posttest at the end of the course on logics. The group of students consisted of 5 women and 2 men aged from 21 to 26.

The students were initially asked about their understanding of the basic concepts in logic. They all revealed a strong understanding of the basic concepts such as validity and soundness of arguments, and distinguished between propositional arguments and syllogistic reasoning with quantifiers. More specifically, they adequately defined syllogistic reasoning as reasoning containing two premises and a conclusion all constituted by one of the four basic forms $a(U,V)$, $e(U,V)$, $i(U,V)$ and $o(U,V)$ with (U,V) being constituted by a subject (S), predicate (P) or middle term (M). The students were not immediately able to reconstruct the conceptual graphs equivalent to the four basic forms in a correct manner.

Following these initial questions the students were presented with the screendump from the test in which they had all participated i.e. with the translations into CGIFs and the derivations presented above.

Looking at the screen the students were asked about a) their ability to read the conceptual graphs, b) their understanding of the rules applied in the example, and c) and their ability to decide the validity of the relevant syllogism. Their answers revealed three interesting trends elaborated on below:

1) The students find the formalism and the deduction rules of conceptual graphs abstract.
2) The students use the natural language formulation of arguments to support them in interpreting the conceptual graphs and not vice versa.
3) The students found the decision procedure to be applied in order to determine the validity of syllogisms to differ markedly from alternative decision procedures such as Venn-diagrams.

Concerning item 1 the students note that the vocabulary of conceptual graphs contains other elements [All: x]”,”[x]”, ”-attr->”, ”If … Then …”, ”~”) than the classic four forms listed above. The students argue that the reading of the conceptual graphs is less intuitive in the sense that it translates differently into natural language than the basic four forms. Thus the *a(S,P)* simply reads ”All S are P” whereas the equivalent conceptual graph

[All: x] [If: (x –attr-> [S]) [Then: (x-attr-> [P])]]

reads “For all x, if x has the attribute S, then x has the attribute P”. The students are not convinced that anything is gained by going to this deeper lever of analysis.

Three students specifically mention the problem of reading conceptual graphs corresponding to the “some”-quantified basic forms, i.e. *i(U,V)* and *o(U,V)*. They find it strange and somewhat surprising that there is no specific conjunction included in the conceptual graph formalism. Seen from a natural language point of view they find that something must be missing in the formulation:

[x](([x] -attr-> [teachers]) ([x] -attr -> [women]))

In general, the students did not find the translation of such formalisms into natural language intuitively convincing.

In the conducted test the students were presented with the syllogism in natural language as well as in the form of conceptual graphs. The purpose of the interview was in part to determine if the conceptual graphs aided the students in their reasoning about the validity of the syllogism. All students noted that they did look at conceptual graphs but that they used the natural language formulation of each of the premises and conclusion in the syllogism to interpret the conceptual graphs and to make the deductions required in order to understand the different formulations of each of the premises and conclusion as conceptual graphs. In short, the natural language formulation of the syllogisms is the fix-point for the reasoning of the students, and they all confirm that they made their evaluations based on the arguments in natural language.

All students found the process of determining the validity of syllogistic arguments by means of conceptual graphs to differ markedly from alternative procedures such as Venn-diagrams. The students univocally state that while Venn-diagrams present a simple mechanic procedure for determining the validity of syllogisms, the process of determining the validity by means of conceptual graphs is markedly different by requiring reformulations of premises and then the ability to determine if the conclusion can be derived by the application of the derivation rules. The students seem to make two points here. First, that the procedure is different in nature than other procedures, and therefore “more difficult” to master. Second, that the procedure involves “more steps” and therefore is less simple than, for example, the use of Venn-diagrams.

The interview reveals that the students find natural language formulations and the classic formulations of syllogisms more intuitive in their evaluation of validity. All students noted, however, that they had received very little training in the formalism of conceptual graphs.

In further studies and experiments on logic teaching new methods should be employed. It may not be satisfactory just to let the students see the formal representations. They should also be allowed to operate on the structures according to the relevant rules. Maybe it would be a good idea to use gamification as suggested in [10]. We may develop a game-like system which allows the students to use the deductive rules, (TR), (SUB), (CON), (MUT) and (EX), on the premises of the syllogism in question in order to decide whether or not the syllogism is provable. The various conceptual graphs may even be presented in a graphical manner and not just in linear form. It might even be possible for the students to draw graphs, which may be significant for a deeper understanding of the structures.

As we have seen, the time needed to evaluate the syllogisms is significantly greater in the post-tests than in the pre-tests. It may be assumed that the use of conceptual graphs gives rise to more reflection and maybe even a better understanding of the notion of validity. An assumption like this may actually be tested performing new statistical tests. In addition to asking the students whether a given syllogism is valid or not, it might be interesting to ask them to evaluate the degree of confidence in the answer given.

6 Conclusions

As we have seen, it is possible to create systems implemented in PROLOG+CG which to some extent can measure the outcome of logic teaching. Such tools can clearly give the logic teacher a lot of important information which will turn out to be useful for the teacher when he or she is planning a logic course. We have also seen that it is possible to create tools which can incorporate the formalism of conceptual graphs (CGIF). However, if such tools should improve the reasoning skills of the students, a new design of the system has to be developed and tested. The qualitative interviews suggest that in order for conceptual graphs to be useful in logic teaching the CG formalism has to be presented more carefully than in the present experiment. An extensive practice will be needed for the students to acquire the skills necessary to read, reason with and evaluate syllogisms in terms of CGs. This will require several lectures - and not just a single lecture. Furthermore, it might be a good idea to develop new and more appealing ways of presenting the conceptual structures. In particular, it might be a good idea to present the structures graphically.

The students have a first-hand experience of the fact that conceptual graphs represent a fundamentally different approach to the evaluation of validity. In fact, it seems that they are aware of the fact that is it straightforward to present deductive aspects in terms of the CG formalism. Consequently, the inclusion of conceptual graphs in the curriculum for courses on logic could add to the understanding of the deductive structures and proof procedures. In order for such teaching to bring this understanding it would – judging from the interview – have to put emphasis on teaching the students a specific decision procedure for determining whether a syllogism can be proved or not. It will not be sufficient - as in the present experiment - just to exemplify how the rules of inference may be used. A more extensive (and if possible) complete proof procedure should be presented. Such a procedure can in fact be developed for the Aristotelian syllogisms in terms of a game-like system, which will allow the user to prove valid syllogisms applying the rules of inference presented above. In this way, the game may in a very practical manner illustrate what it means prove a theorem in a deductive system. In addition, the game may in a very clear manner illustrate that 15 of the valid syllogisms can be proved without the use of the (EX) rule, whereas the proofs of 9 of the syllogisms which are valid from an Aristotelian point of view will have to involve the use of the (EX) rule along with the other rules of inference mentioned above.

References

1. Aristotle: Prior Analytics, Jenkinson A.J. (transl.). The Internet Classics Archive (1994-2000), http://classics.mit.edu//Aristotle/prior.html
2. Kabbaj, A., Janta-Polczynski, M.: From PROLOG++ to PROLOG+CG: A CG object-oriented logic programming language. In: Ganter, B., Mineau, G.W. (eds.) ICCS 2000. LNCS (LNAI), vol. 1867, pp. 540–554. Springer, Heidelberg (2000)
3. Kabbaj, A., Moulin, B., Gancet, J., Nadeau, D., Rouleau, O.: Uses, improvements, and extensions of Prolog+CG: Case studies. In: Delugach, H.S., Stumme, G. (eds.) ICCS 2001. LNCS (LNAI), vol. 2120, pp. 346–359. Springer, Heidelberg (2001)
4. Parry, W.T., Hacker, E.A.: Aristotelian Logic. State University of New York Press (1991)
5. Petersen, U., Schärfe, H., Øhrstrøm, P.: Online Course in Knowledge Representation using Conceptual Graphs, Aalborg University (2005), http://cg.huminf.aau.dk
6. Petersen, U.: Prolog+CG: A Maintainer's Perspective. In: de Moor, A., Polovina, S., Delugach, H. (eds.) Proceedings of the First Conceptual Structures Interoperability Workshop (CS-TIW 2006). Aalborg University Press (2006)
7. Sowa, J.F.: Conceptual Graph Summary, http://www.jfsowa.com/cg/cgif.htm (accessed February 2014)
8. Øhrstrøm, P., Sandborg-Petersen, U., Ploug, T.: Syllog - A Tool for Logic Teaching. In: Proceedings of Artificial Intelligence Workshops 2010 (AIW 2010), Mimos Berhad, pp. 42–55 (2010)
9. Øhrstrøm, P., Sandborg-Petersen, U., Thorvaldsen, S., Ploug, T.: Classical Syllogisms in Logic Teaching. In: Pfeiffer, H.D., Ignatov, D.I., Poelmans, J., Gadiraju, N. (eds.) ICCS 2013. LNCS (LNAI), vol. 7735, pp. 31–43. Springer, Heidelberg (2013)
10. 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.) EC-TEL 2013. LNCS, vol. 8095, pp. 410–423. Springer, Heidelberg (2013)