Syllogism.studio: Visualising Syllogisms for Inf1A

Sioned Jones

4th Year Project Report
Computer Science
School of Informatics
University of Edinburgh

2021

Abstract

This report documents the design and evaluation of Syllogism Studio, a web application catered to informatics students studying syllogisms. The application focuses on visualising syllogisms through Venn diagram representations and supports students in using these representations as reasoning tools. The application is then utilised to conduct a study investigating how Venn diagram representations of syllogisms affect syllogistic reasoning.

Acknowledgements

First and foremost, I must thank my supervisor, Professor Michael Fourman, for introducing me to Syllogisms in our first meeting. While this introduction led to months of confusion as I learned to navigate the world of validity and Venn diagrams, the ongoing support provided meant that this project remained intellectually challenging, but not overwhelming. Without your guidance, I would still be confused.

Secondly, I must thank Sarah, Peter, and Pravar, for the weekly meetings that kept me motivated. Without your contributions, I would be uninspired.

Finally, thank you to Uplands Roast for the 58 coffees that fuelled this project. Without your coffee, I would be asleep.

Table of Contents

Chapter 1

Introduction

This report covers the design, creation, and evaluation of a web application to support learning in Informatics 1A - Introduction to Computation. The course is aimed at first-year informatics students, serving as an introduction to programming, computation, and logic.

The computation and logic strand of the course, which is the focus of this project, covers five main topics: syllogisms, inference and reduction, karnaugh maps, satisfiability, and finite state machines.

For the first time this year, students were assessed with a test on each topic during the semester. From a research standpoint, this was most convenient, as it provided a breakdown of marks by topic.

The semester's test results highlighted the first topic covered, syllogisms, as the area where students were most likely to under-perform. The reasons for lower marks on this topic are unclear. It could be due to this test being the first university-level assessment for many, teething problems with the new assessment method, or simply that syllogisms are a particularly tricky topic. Regardless, there is a strong case for supplementary material to better support students learning syllogisms as part of Inf1A.

Before jumping right into the proposed solution to the problem, the next chapters summarise the ideas that informed this project. Chapter 2 provides a broad overview of syllogisms, non-logical factors that lead to errors in syllogistic reasoning, and how these ideas link to the Inf1A course content. Chapter 3 discusses a selection of pre-existing tools for learning syllogisms, that are evaluated in relation to the needs of Inf1A students.

The next few chapters move on to the more practical component of the project. Chapter 4 concerns the design and implementation of Syllogism Studio, a tool to assist students in learning syllogisms. Chapters 5 and 6 discuss user studies conducted to evaluate the tool.

Chapter 7 reflects on the project and discusses what further work can be done.

Chapter 2

From Aristotle to Inf1A

This chapter lays the groundwork for the project, by summarising the background research conducted. The ideas discussed in this chapter will later be used to evaluate current tools for learning syllogisms, and inform the design of my own.

2.1 An introduction to syllogisms

A syllogism is a form of deductive reasoning first described by Aristotle around 350 BCE in his work Prior Analytics. For those interested in discovering syllogism through Aristotle's own words, the English translation of the text [9] provides a thorough explanation of syllogisms. For those wanting an introduction to the key points, consider the following:

> If I work hard on this project, I will create a helpful tool for learning syllogisms. If I create a helpful tool for learning syllogisms, the project will be a success. Therefore, if I work hard on this project, the project will be a success.

The above inference is an example of how a student might motivate themselves to work on their Honours project. It is also an example of a syllogism.

As is typical for a syllogism, the above example consists of two premises, namely "If I work hard on this project, I will create a helpful tool for learning syllogisms" and "If I create a helpful tool for learning syllogisms, the project will be a success". From the premises of the argument, a conclusion is inferred, in this case, "if I work hard on this project, the project will be a success". The above syllogism is considered sound, as the conclusion follows logically from the premise. That is, in any universe where the premises are valid, the conclusion is also valid.

The two premises of this syllogism describe the relationship between the terms "working hard on a project", "creation of a helpful tool" and "a successful project". However, syllogisms are not confined to discussing Honours projects. Any terms can be substituted without affecting the soundness of the argument. For example, one could argue:

All doctors are tall
All aliens are doctors
Therefore, all aliens are tall

In the natural universe, the argument is absurd. It seems incredibly unlikely that there are aliens that are doctors, and that there does not exist a doctor that is not tall. However, in a universe where the premises are valid, it would indeed be the case that all aliens are tall, rendering the syllogism sound. Note that the soundness of a syllogism is defined by the validity of the conclusion when the premises are valid. As explained by Copi and Cohen [14] a conclusion that does not follow logically from the premises makes for an unsound argument, regardless of its validity in the natural universe. The previous example demonstrates the inverse of this statement. That is, a conclusion that would be invalid in the natural universe, but follows logically from the premises makes a sound argument.

As the exact terms described by a syllogism do not affect its soundness, it is often simpler to represent syllogisms in their standard form, where the terms are replaced by the letters S, M, and P. Here, S represents the subject, or the first term, of the conclusion. P represents the predicate of the conclusion, which is the second term that makes an assertion about the subject, S. Finally, we have M, the middle term that links S and P together, by acting as a subject or predicate in each premise.

With standardised terms, the two premises can be distinguished. The premise that contains both M and P is called the major premise, then the premise linking M and S the minor premise.

In the example, both premises and the conclusion are of the form "all A are B", which is described as universal affirmative. However, a statement can be universal or particular, affirmative or negative, giving rise to 4 possibilities, abbreviated as A, I, E, and O: [1]

Universal affirmative: All A are B, abbreviated as A
Particular affirmative: Some A are B, abbreviated as I
Universal negative: No A are B, abbreviated as E
Particular negative: Some A are not B, abbreviated as O

The quality (affirmative/negative) and quantity (universal/particular) of the statements are described as the mood [13]. With 3 statements per syllogism and 4 combinations of their quality and quantity, there are a total of (4^3) 64 possible moods for a syllogism.

Then consider that the order of the terms within the two premises can differ, giving rise to a syllogism's figure, which is dependent on the middle term's placement within the premises. The middle term M can be the subject or the predicate of each premise, resulting in (2^2) 4 figures, summarised in table 2.1.

64 possible moods, and 4 possible figures, results in a total of (4×64) 256 syllogisms.[2]

[1] A and I are derived from the Latin "affirmo" meaning to affirm, and E and O from "Nego", meaning to deny [14].

[2] It's been argued that there are 512 syllogisms, as the minor premise could come before, or after the major premise [22]. However, for this project syllogisms are only presented in their standard form, with the major premise proceeding the minor premise, creating a problem space of 256 syllogisms.

$M-P$	$P-M$	$M-P$	$P-M$
$S-M$	$S-M$	$M-S$	$M-S$
First Figure	Second Figure	Third Figure	Fourth Figure

Table 2.1: Syllogistic Figures

Of these 256 syllogisms, only 24 are considered sound, with a conclusion that follows from the premises. For the 232 unsound syllogisms, the conclusion does not follow from the premise. An example of such a syllogism is:

No mothers are fathers
No fathers are athletes
Therefore, no mothers are athletes

In the above argument, the conclusion cannot be inferred from the premises. In a universe where the premises are valid, it could be the case that no mothers are athletes, however, it could also be the case that some, or all mothers are athletes. When the premises do not provide enough information to accept the conclusion with certainty, the syllogism is unsound.

The 24 sound syllogisms are further categorised as conditionally, or unconditionally, sound. For unconditionally sound syllogisms, the conclusion is always valid in a universe where both premises are valid. Conditionally sound syllogisms only hold when a specific term exists, by having at least one instance.

Going back to a previous example:

All doctors are tall
All aliens are doctors
Therefore, all aliens are tall

The conclusion "all aliens are tall" is equivalent to "no aliens are not tall" which should be valid regardless of whether there are any aliens in the universe. However, if there are indeed aliens, (that is, "aliens" has at least one instance) it would also be the case that "some aliens are tall", giving rise to another sound syllogism. The existence of aliens, or the S term, is the condition that needs to be met for the following syllogism to be sound:

All doctors are tall
All aliens are doctors
Therefore, some aliens are tall

A table of conditionally and unconditionally sound syllogisms is provided in Appendix A.

2.2 Errors in Syllogistic Reasoning

With only 24 of the 256 syllogisms having a conclusion that follows from the premises, how can the few sound syllogisms be distinguished?

One option is to memorise all 24 sound syllogisms. However, if the purpose of studying syllogisms is to improve one's reasoning skills, this defeats the object of the practice. A person may be able to identify a sound standard syllogism, but the method fails when extending to multi-premise, conditional or disjunctive syllogisms.

A better option would be to take one sound syllogism from each figure to act as an axiom. The remaining sound syllogisms can then be derived by contraposition and substitution.

Again, is this necessary? As explained in Bradley's Logic [27], if we need axioms to reason, how did we reason before axioms?

Although axioms can be a helpful tool, they are not essential to reasoning. People of all ages, children included, have an ability to reason with syllogisms without any training [10]. This is unsurprising, as reasoning is something we practice daily, often without realising [16]. However, the ability to reason is not the ability to reason without error, and there are many opportunities to make errors with syllogisms.

Atmosphere

Recall from the previous section the concept of the mood of a syllogism, and the abbreviations A, I, E, and O for the form of each statement. The atmosphere effect describes the tendency to accept a conclusion whose form matches the form of the premises [34]. Where the form of the two premises is the same, they create a bias towards a conclusion of the same form. Where the form of the premises differs, a negative premise has the strongest effect, followed by that of a particular premise.

For example, take the following premise pair, from which no valid conclusion can be drawn:

All M are P
Some S are not M

Under the atmosphere effect, a reasoner is biased to conclude that "Some S are not P", due to the atmosphere created by the negative particular premise, even though it is not a valid conclusion.

A re-examination of the supporting data published by Woodworth and Sells [34] revealed that not all participants were subject to the atmosphere effect [33]. Participants that appeared to use logic to determine the validity of a conclusion were not biased by the atmosphere effect, regardless of the accuracy of their logic. However, the remaining participants are believed to use the atmosphere of the premise to decide on the validity of a conclusion, suggesting that atmosphere is used to inform guessing.

Belief Bias

Another non-logical factor that influences reasoning is personal belief [26], as people are more likely to accept a conclusion if it aligns with their worldview. This effect is seen the most when faced with an invalid conclusion [17], suggesting it acts as a method of determining validity where the participant has failed to derive the conclusion through logical reasoning.

Belief biases highlight how the terms of a syllogism affect its perceived soundness. For example, take the two conclusions “Some women are not people” and “Some people are not women”, both of the form “Some A are not B”. A belief bias would make it more likely for the first example to be deemed invalid, as it goes against a subject’s understanding of the world. Similarly, “Some people are not women” would be more likely to be accepted under the belief bias, as it is observed, and therefore believed, to be valid. As explained by Woodworth [35], being certain of the truth or falsity of a conclusion distorts the ability to reason impartially.

Emotion

Beliefs that invoke emotion further distort the ability to reason. A study conducted in 1941 [23] provided evidence towards emotion’s effect by presenting syllogisms with conclusions that were highly controversial and emotionally charged, especially given the political climate at the time.

As expected, there was a normal distribution for the accuracy of validity judgements on unemotional syllogisms. However, for the syllogisms covering topics such as war, communism, and antisemitism, there was a J-shaped distribution, with a large number of participants making few correct judgements. The study illustrates how emotion can conflict with reasoning, which makes it difficult to reason objectively with content that is emotionally charged [16].

Language Ambiguity

Consider the statement “All athletes are on the running track”. Does this imply that there are athletes on the running track? In everyday speech, it would be reasonable to assume that there are some athletes, all of which are on the running track. However, logically, this does not follow. Take the case where there are no athletes. The statement that all athletes are on the running track, or its equivalent “no athletes are not on the running track” is valid, but the earlier assumption that there are some athletes on the track is not.

Next, take the statement “some of the girls went to the shop”. Where are the rest of the girls? Naturally, we’d assume that some of the girls did not go to the shop. However, this conclusion does not follow logically. It could be the case that all of the girls went to the shop. In logic, “some” keeps to its definition of an unknown quantity of at least one. In other words, some means “some, and possibly all” rather than “some, but not all”.

The difference in how quantifiers are used in everyday speech versus in logic can be a source of confusion. In natural language, "all" and "some" have additional implied meaning in many contexts, and this learned implication can be unwittingly carried on into logic, leading to the construction of invalid conclusions.

These four non-logical sources of error are born from the language used to express syllogisms, which fails to express the true relationship between terms [35]. Language ambiguity's link to language needs no explanation. With emotion and belief bias, the terms used in the syllogisms, and their associations derail one's ability to reason.

Explaining the atmosphere effect as a natural language problem is not as simple, as this effect is observed in reasoners without prior training in logic. It's likely that removing natural language to express syllogisms in symbolic notation would reduce the atmosphere effect, but not necessarily because the effect comes from natural language itself. Rather, to understand a different notation, there is a need for some understanding and training in logic. Participants with this training are more likely to be able to reason logically to reach a conclusion, rather than form a guess influenced by atmosphere.

Despite the uncertainty regarding whether the atmosphere effect is an artifact of natural language or not, the remainder of this project builds on the hypothesis that the atmosphere effect could be reduced, at least in part, by removing natural language. Chapter 5 discusses an attempt to prove, or disprove this hypothesis.

2.3 Notation and graphical representations

While it is possible to explore syllogisms in natural language, this method has been shown to be error-prone [21] [31] as the factors outlined in the previous section come into play, influencing one's ability to reason. One solution to the natural language problem is to use symbolic notation, mapping terms to letters, and using the double turnstile symbol to represent relationships between the terms, as seen below.

$A \vDash B$	$A \nvDash \neg B$	$A \vDash \neg B$	$A \nvDash B$
All A are B	Some A are B	No A are B	Some A are not B

Table 2.2: Premises in Sequent Form

For those familiar with sequent form, this removes ambiguity in meaning. Another benefit is that terms are mapped to letters, so reasoners can't be biased by the terms used.

However, this notation can make syllogisms less accessible, as a reasoner must learn the language of logic before reasoning with it. For those unfamiliar with the notation, more effort could be focused on interpreting the language than on reasoning itself. Because of this, there can be a temptation to simply translate the notation into natural language, reason verbally, then translate back into sequent form. This not only defies the purpose of using the notation in the first place, but opens new pathways for errors in the translation.

While using symbolic notation has benefits, including a level of abstraction from the subject, and the conciseness of the notation, its use could do more harm than good for those new to logic. This is especially true for Inf1A students, whose focus is on developing a logical way of thinking, rather than learning the formalities of logic.

A solution to this problem is to use visual representations, which can support reasoning whilst being more accessible [12]. Venn diagrams, for example, are widely used outside logic, so a student learning to express syllogisms with Venn diagrams can adapt their existing knowledge, rather than learn a new notation.

In a Venn diagram representation, the predicates of a proposition are represented by overlapping circles, and the relationships between them expressed through denoting inhabited and uninhabited regions, as follows:

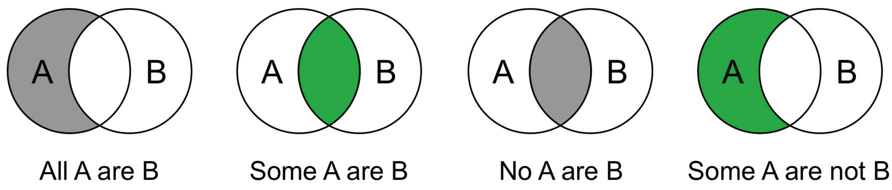

Figure 2.1: Venn diagram representation of propositions. An uninhabited region is shown in grey, and an inhabited region in green. White regions may, or may not, be inhabited.

These individual diagrams can then be combined into a triple Venn diagram to represent a syllogism, shown in figure 2.2.

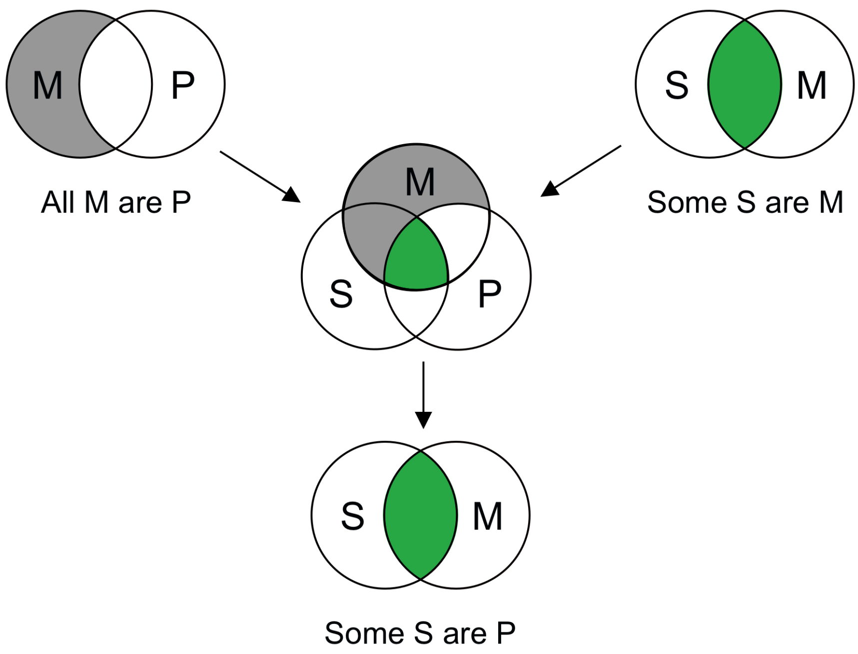

Figure 2.2: Example of how two premises can be combined on a triple Venn diagram, and used to draw a conclusion.

When teaching and learning syllogisms, visual representations bring the additional benefit of showing how the reasoner reached their conclusion.

Consider a scenario where an Inf1A student has incorrectly judged the soundness of a syllogism, or derived an invalid conclusion. When reasoning verbally, there is little scope for an instructor to identify where the error is made, unless the reasoner provides a written explanation of how they reached their conclusion. The reasoning process is essentially a black box. There is an input and output, but the reasoning process itself remains a mystery. When using visual representations to reason, the students externalise and illustrate their thought process, and it becomes possible to identify the particular step at which their method failed.

While there is a benefit from an instructor's perspective to using visual representations, visual methods also benefit the students. Once a student has represented the premises of a syllogism as a Venn diagram, it is much simpler to find, or validate, a conclusion. Instead of reasoning through all the possible conclusions, the problem can now be solved by simply observing the diagram. Which regions of S and P are marked as inhabited? Which regions are uninhabited? And finally, what does this mean for the relationship between S and P? What was a philosophical problem has been abstracted, and simplified to purely an observational one.

Venn diagram representations for syllogisms bring the benefits of abstraction, externalising the reasoning process, and visualising a student's thought process, often without having to learn a new language or visual representation. However, a new source of error is created - the translation of statements to, and from, their linguistic and diagrammatic representations. How can the potential for error in translation be mitigated? This project proposes a tool to do precisely that.

2.4 Syllogisms in Inf1A

The Inf1A course content on syllogisms aligns with what has been covered in this chapter. Students are expected to be able to reason with categorical syllogisms, convert these syllogisms to standard form, and use Venn diagrams to assist with and illustrate their reasoning. Understanding the existential assumption is also essential, as it is a common theme in exam questions.

The challenge of this project is to create a tool that allows students to explore these concepts, whilst preventing them from making common errors in reasoning. This assistance needs to be offered in a way that meaningfully contributes towards the learning outcomes of the course, rather than providing shortcuts to solutions.

With Inf1A's recent move to an entirely online learning environment, the tool should be presented as a web application. As it's too early to say if and when in-person teaching will resume, a solution should take advantage of the online environment, but also be designed to outlive it. That is, be appropriate for use by a single student studying online, by a group in a tutorial, or even by a lecturer to illustrate concepts.

2.5 Chapter Summary

This chapter has, as the chapter title suggests, taken us all the way from Aristotle to this year's Inf1A course. Having covered over 2,000 years of syllogisms in 8 pages, let's take a recap.

What are syllogisms?
Syllogisms are a form of logical argument, where a reasoner takes some information (the premises) and infers more information (the conclusion).

What problems do people face when learning syllogisms?
Four key non-logical factors can mislead people when reasoning with syllogisms: atmosphere, belief, emotion, and language ambiguity. These factors can be especially prominent in people that are learning syllogisms for the first time, such as Inf1A students.

How can these issues be prevented?
This is a key conceptual challenge. I propose using a non-linguistic notation for syllogisms as a potential solution.

What do you mean by "non-linguistic notation"?
It could mean symbolic notation using the double-turnstile, or any one of the visual representations people use. Specifically for this project, we focus on Venn diagram notation, as it benefits both instructors and learners in Inf1A.

How does all of this link to Inf1A?
The ideas in this chapter are used to inform the design of a tool to help Inf1A students learn syllogisms.

Chapter 3

Existing Tools

This chapter provides a critical evaluation of tools already available for learning Syllogisms and generating Venn diagrams. My own design will later adapt and incorporate successful aspects of these tools, whilst implementing solutions to their shortcomings.

3.1 Syllogism Quizzes

There are several tools available that present syllogisms in a quiz format, where users can assess the soundness of a syllogism. See Philosophy Experiment's quiz [4], Fibonicci's Syllogisms Aptitude Test [3], or Tutorials Point's online test [2] as examples. Such tools are insufficient for learning syllogisms for the same reason; they do not go beyond a limited number of pre-defined questions. Once the quiz is complete, the tool is obsolete, and cannot continuously support students over the semester.

One quiz-style tool that does not have this drawback is Syllog [24], a web application designed to measure students' ability to reason with syllogisms [30]. Syllog generates syllogisms from a database of terms, forms, and figures, rather than presenting hard-coded questions [29]. For this reason, Syllog is more appropriate as a practice tool as students will not run out of questions to answer.

A study that used Syllog as a way of assessing syllogistic reasoning in Philosophy and Informatics students [28] suggests that it does not cater well to Informatics students' needs. Philosophy and informatics undergraduates used Syllog before, and after, taking a course in logic. The philosophy students outperformed informatics students in the pre-test then went on to achieve even better scores in the post-test. For informatics students, there was no evidence of similar improvements.

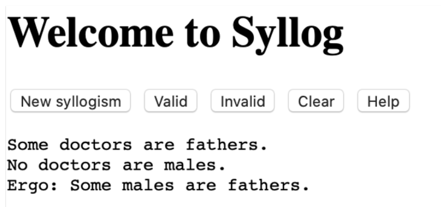

Figure 3.1: Sample Syllog question [24]

The researchers suggest that the philosophy students could have had more interest in the course, and learned more. However, it seems unlikely that the informatics students did not learn from the course. An alternative explanation is that Syllog does not cater well to informatics students' way of thinking, and so their learning did not translate into higher Syllog scores. As the syllogisms are presented in natural language, this could cater more to philosophy students, who typically express their arguments verbally. For informatics students that are more accustomed to expressing ideas in proofs, counterexamples, and diagrams, would a tool that presents syllogisms diagrammatically be more appropriate? If Syllog catered specifically to informatics students' way of thinking, the pre and post-test scores may have captured informatics students' improved understanding after the course.

3.2 Visual Tools for Informatics

Since syllogism quiz tools do not meet the needs of informatics students, we'll now focus on tools that are designed by, and for, The School of Informatics.

In previous iterations of this project, only one tool has been created for syllogisms, as the topic was only added to Inf1A in 2019. This tool is Soak up Syllogisms [7], designed as a teaching tool to help users understand syllogisms [8].

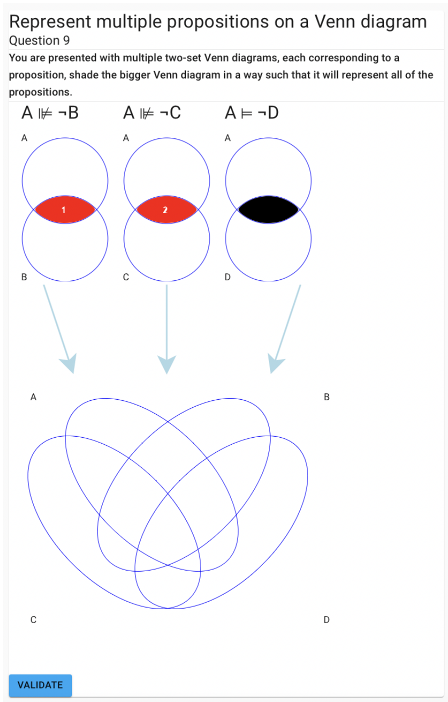

Figure 3.2: Sample question from Soak up Syllogisms [8]

Having been designed specifically for Inf1A, Soak up Syllogisms succeeds in covering the breadth of knowledge expected of students. The web application offers practice questions on creating and interpreting Venn diagrams, as well as translation to and from different notations. Unfortunately, the application covers *too much*. There are questions with more than the standard two premises and three terms, which could confuse and overwhelm students (see figure 3.2). As the application offers a limited number of questions, every question beyond the course content would be better used to cover key knowledge that students need.

Soak up Syllogisms also has an instant solver, which takes a syllogism and determines whether it is sound. Its application to any Syllogism a student could come across makes this a valuable feature. Despite this feature's potential, the interface makes it too easy to make errors, as each term must be retyped every time it is used. Even with a 2 premise syllogism, there are 6 opportunities for error. When an error is made, the error message is unhelpful. See figure 3.3, where a better error message would highlight where the error is made, for example "Error: A and a are only used once". While it's reasonable to assume that a user could find their error easily with 2 premises, it would get much more difficult as the number of errors and premises increase.

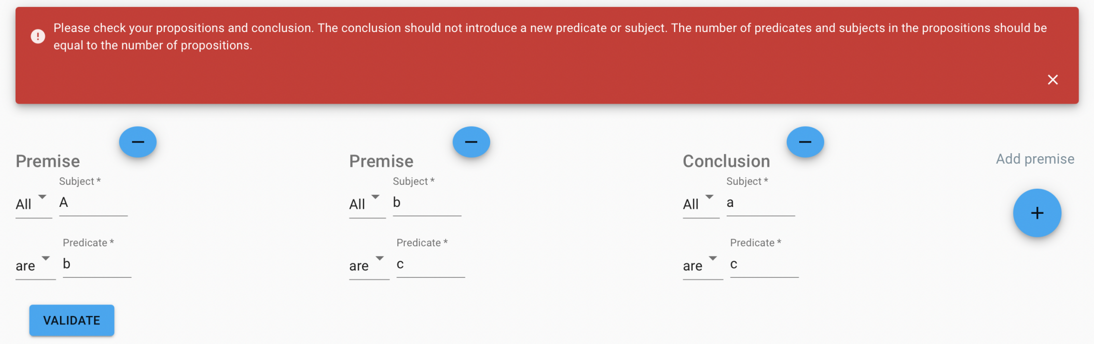

Figure 3.3: Input error, and corresponding error message from Soak up Syllogisms' instant solver [8]

While better error messages would be helpful, the ideal solution is an interface that prevents these errors from being made. An example of such an interface is Little's Syllogism Validity Checker [25], pictured in figure 3.4. Here, a user must only type each term once, preventing the error shown in figure 3.3. However, Little's interface allows some errors to go unnoticed, such as if a student inputted "All A are not A". Another drawback is that it does not allow input of the form "No A are B", requiring the user to input the equivalent "All A are not B". Solutions to these issues are implemented in this project and will be discussed in Chapter 4.

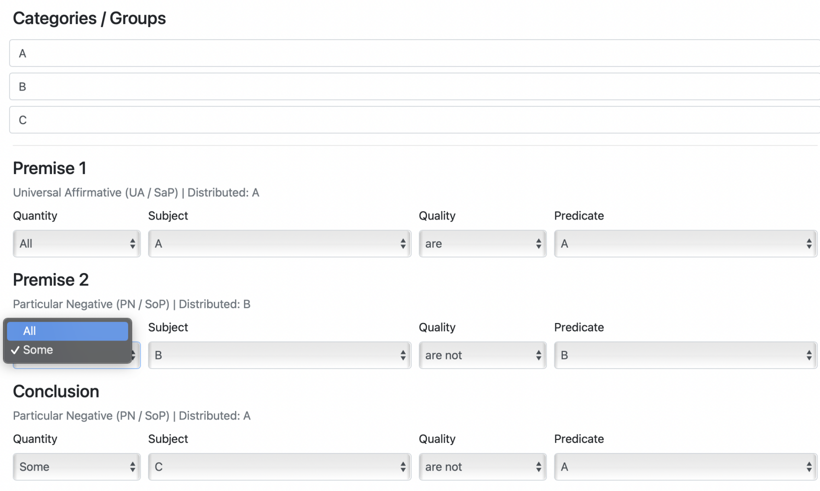

Figure 3.4: Syllogism input form from Little's Syllogism Validity Checker [25]

Another tool designed specifically for Inf1A is the Venn Diagram Generator [6], which takes a Boolean expression written in JavaScript notation and outputs the corresponding Venn diagram. Along with the relevant Venn diagram, the tool outputs another 15 diagrams, shown in figure 3.5. The reason for these additional, largely irrelevant diagrams is unclear. Unless there's a specific reason for their inclusion, the additional diagrams should not be displayed by default, as they are likely to confuse users and distract from the purpose of the tool.

Addressing this issue is a student-made adaptation of the Venn Diagram Generator [5],

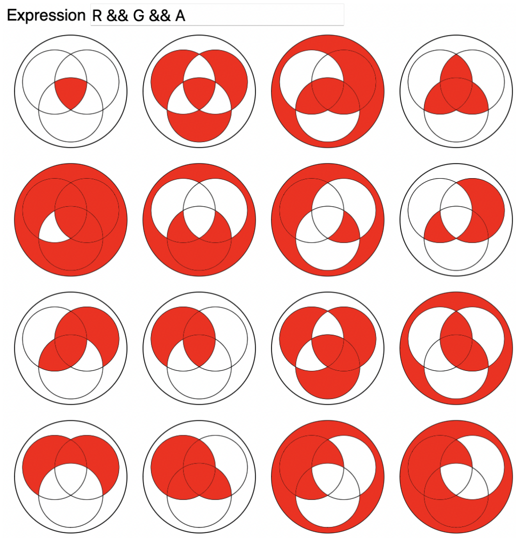

Figure 3.5: Venn diagrams outputted by the Venn Diagram Generator [6]

which only displays the one, relevant, Venn diagram, as seen in figure 3.6. This version of the generator also improves on the previous by allowing a wider range of notation in the text input box and providing a clearer guide to what input is valid. However, neither the original nor adapted version can take two premises with different quantities as input, making them inapplicable to the study of syllogisms. This is because the regions in the Venn diagram can only have two states, white or red for true or false. For syllogisms with premises of different quantities, three states need to be displayed simultaneously: inhabited (true), uninhabited (false), and unknown. As it can only deal with a small number of possible syllogisms, the generator is insufficient for displaying syllogisms.

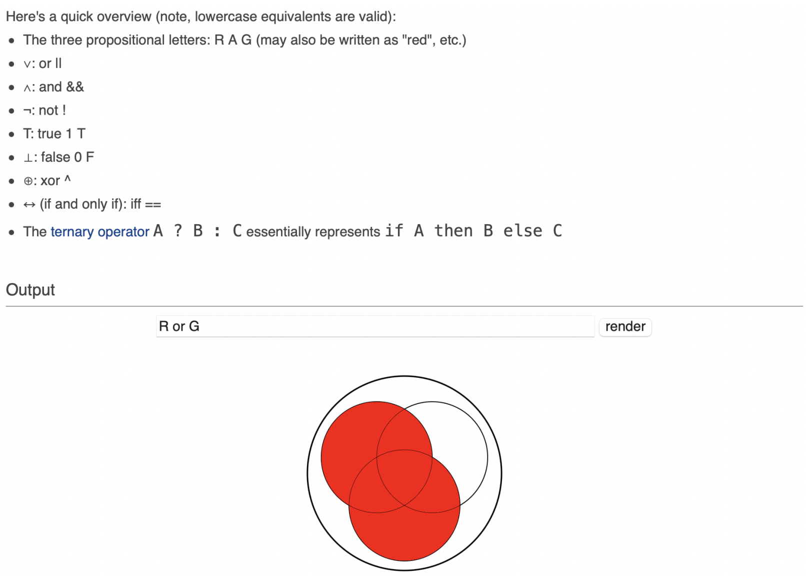

Figure 3.6: Adapted version of the Venn Diagram Generator [5]

3.3 Chapter Summary

Chapter 3 had taken us through the syllogism tools that are already available online, how they could apply to Inf1A, and how they fall short.

Why are you making a new web app for syllogisms, when there are already so many out there?

There are many pre-existing tools, yes, but they tend to have a limited number of questions, so a student can't keep using these tools over the semester.

What if one of these tools had so many questions that they would never run out?

This is better. However, I've yet to find a tool that does this whilst catering to the needs of Inf1A students specifically.

What about the tools that have already been created for Inf1A?

There is a validity checker available, as well as a Venn diagram generator. What's needed is a tool that helps students determine the soundness of a syllogism themselves, with the help of a Venn diagram generator that can represent syllogisms.

Chapter 4

Design and Implementation

This chapter covers the design and implementation of Syllogism Studio. The design process followed an iterative and incremental approach, beginning with designing and implementing a minimum viable product, then adapting and extending the app as necessary. To allow this adaptive approach to design and implementation, the design process was guided by a set of goals rather than strict requirements.

4.1 Goals

The goals for design and implementation are all drawn from the research covered in Chapters 1-3. The chapter numbers in brackets correspond to the chapters that motivate each goal.

1. To support students in learning syllogism (Chapter 1)
2. To decrease the likelihood of non-logical factors influencing students' reasoning (Chapter 2.2)
3. To assist students in using visual representations (Chapter 2.3)
4. To increase understanding, rather than providing a shortcut to solutions (Chapter 2.4)
5. To be usable in different social contexts (Chapter 2.4)
6. To generate new syllogisms for students to explore (Chapter 3.1)
7. To cater specifically to the needs of informatics students studying Syllogisms (Chapter 3.1)
8. To prevent errors when inputting syllogisms (Chapter 3.2)
9. To focus students' attention on the key information (Chapter 3.2)

4.2 Early Version

The initial implementation of Syllogism Studio focused on how premises could be inputted and presented visually. The input was based on the interface of Little's Syllogism Validity Checker [25], allowing students to input premises as they are presented in the course, without the need to translate into a different notation. Little's input form was adapted in the following ways:

1. Added "No" to the drop-down list so that students wouldn't need to translate a premise such as "No A is B" to "All A is not B" before inputting, as this could be a source of error.

2. Required the terms to be defined as S, M, and P, rather than inputting the names of the terms in any order. This was done so that students can practise identifying S, M, and P.

3. Labelled the premises as "Major Premise" and "Minor Premise" so that students would understand the difference between the two premises. The "S" term cannot be inputted into the major premise, and the "P" term cannot be used in the minor premise.

4. Added an auto-complete feature to the premise inputs. If M is inputted as the subject of the major premise, the predicate is automatically selected as P. This feature is both convenient for the user, and prevents errors, by automatically correcting illegal input such as "All M are M".

5. Removed unnecessary features such as conclusion input, validity tests, and information on the distribution of terms.

The visual representation rendered by Syllogism Studio was inspired by the Venn diagram generators discussed in Chapter 3.2. The idea was to take the premises and simply generate a Venn diagram representation, as shown in figure 4.1.

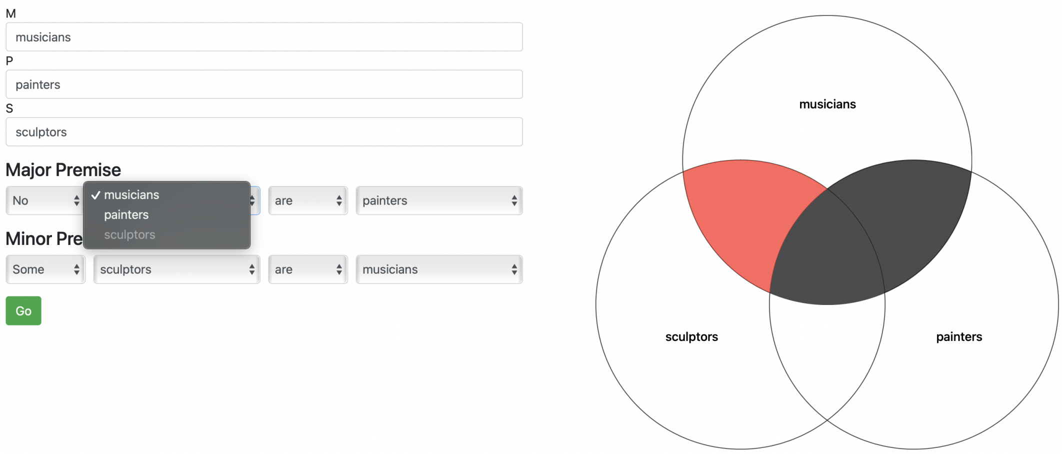

Figure 4.1: Initial interface of Syllogism Studio. Red denotes inhabited, and grey uninhabited, areas.

Initially, D3.js [11] seemed like the most appropriate tool for the job, as it would allow Venn diagrams to be generated. However, after learning to use D3 to draw Venn diagram SVGs, I found a much simpler solution - to draw the SVGs myself using a vector graphic editor. Although using D3, or drawing a Venn diagram using a vector graphic editor would produce the same result, I decided to stick to the simpler method to make it easier for future users to adapt and extend my work without requiring knowledge of D3.

The translation from premise to Venn diagram was done in relatively few, simple steps. First, the terms are mapped back to their standard letters. Then the form of each premise is abbreviated as either A, I, E, or O. Each premise is now reduced to three characters, for example, a universal negative premise with M as the subject and P as the predicate would be stored as MeP. Finally, this three-character string is displayed by changing the fill colour of the relevant SVG elements.

This early version succeeded in achieving, or partially achieving, a number of the design goals, with as simple an implementation as possible. A summary of which goals were achieved, and which ones remained to be satisfied is provided in table 4.1.

Table 4.1: Evaluation against goals

Goal no.	Achieved	Description	Explanation
1	inconclusive	support students in learning Syllogism	a user study should be conducted with Inf1A students to assess whether they find the tool useful
2	inconclusive	decrease the likelihood of non-logical factors influencing student's reasoning	a study should be conducted to investigate Syllogism Studio's effect on reasoning
3	partially	assist students in using visual representations	generates Venn diagrams, but does not give explanation or guidance on how to interpret them
4	no	increase understanding, rather than providing a shortcut to solutions	provides a shortcut to a correct Venn diagram without explanation
5	yes	be usable in different social contexts	instructors could use in class to quickly generate Venn diagrams, and individual student could use to help with creating or validating diagrams
6	no	generate new syllogisms as needed	no question generator is implemented
7	partially	cater specifically to the needs of informatics students studying syllogisms	assists with standard form and Venn diagrams, but doesn't provide practice questions, or validate students' conclusions
8	yes	prevent errors when inputting syllogisms	input form prevents errors
9	no	focus students' attention on the key information	all the available information is presented on one page

4.3 Extended Version

The next challenge was to extend the application to satisfy as many of the goals as possible, without over-complicating the design.

To satisfy goals three and four, a guide to interpreting the Venn diagrams needed to be added in a way that would not clutter the interface nor confuse the user. The inspiration for my solution came from The Pudding [1], a digital publication that presents complex, debated topics visually. The visuals in their essays speak for themselves, with minimal annotation to enhance understanding. To achieve a similar result in Syllogism Studio, I broke the creation of the Venn diagram down into smaller steps and followed The Pudding's guide to scrollama.js [19] [18] to add a simple explanation to each step of its creation, such as in figure 4.2. This provided a guide on both how to draw, and interpret the Venn diagram at each step, helping users to understand the process of using visual representations, rather than jumping straight to the final product.

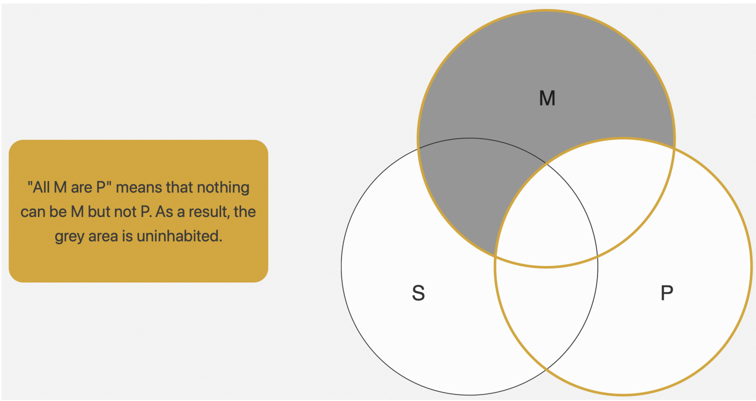

Figure 4.2: A Venn diagram presented by Syllogism Studio

A by-product of this redesign is that it satisfied goal nine. As users scroll through the interface, only the most relevant information is presented at any given time. Initially, the user is presented with the input form. Later, only the Venn diagram is presented along with a brief explanation of what it represents. To further focus the user's attention on the key information, the areas of the Venn diagram that are being discussed at any particular step are highlighted with a thick, gold outline, visible in figure 4.2.

A drawback of this adaptation is that it made the application less usable as a teaching tool in lectures and tutorials. It would be time-consuming for an instructor to scroll through the interface to display a single Venn diagram to students. The idea of adding a "Skip" button was explored but eventually dropped, as it would encourage students to skip to a solution rather than working through a problem. Here, there was a trade-off between goals, and it was decided to create a tool designed for an individual student, rather than catering to lecturers and tutors too.

After explaining the Venn diagram, an additional text box was added to challenge users to draw a conclusion from the Venn diagram, shown in figure 4.3. As the premises were inputted in standard form, it was possible to reduce the possible conclusions to 5 options; All S are P, Some S are P, No S are P, Some S are not P, and None of the above. Students can enter any two premises, and test themselves on the conclusion, meaning the app does not fall into the trap found in many syllogism quizzes of only having a certain number of questions to answer. However, asking users to define their own questions is not a perfect solution, so a randomize button was added to the input form. This randomly selects a quantifier, qualifier, and figure for the premise, before users work through the app to find a conclusion. This satisfies goal six, by generating new syllogisms for the students to practise.

To fully address goal seven, whenever a premise pair that have a conditionally valid conclusion is inputted or generated, and additional dialogue is presented. Here, students are asked to draw an additional conclusion based on one of the regions being known to be inhabited, as shown in figure 4.3. This closely follows the structure of many exam or test questions in Inf1A, and so caters specifically to the needs of Inf1A students.

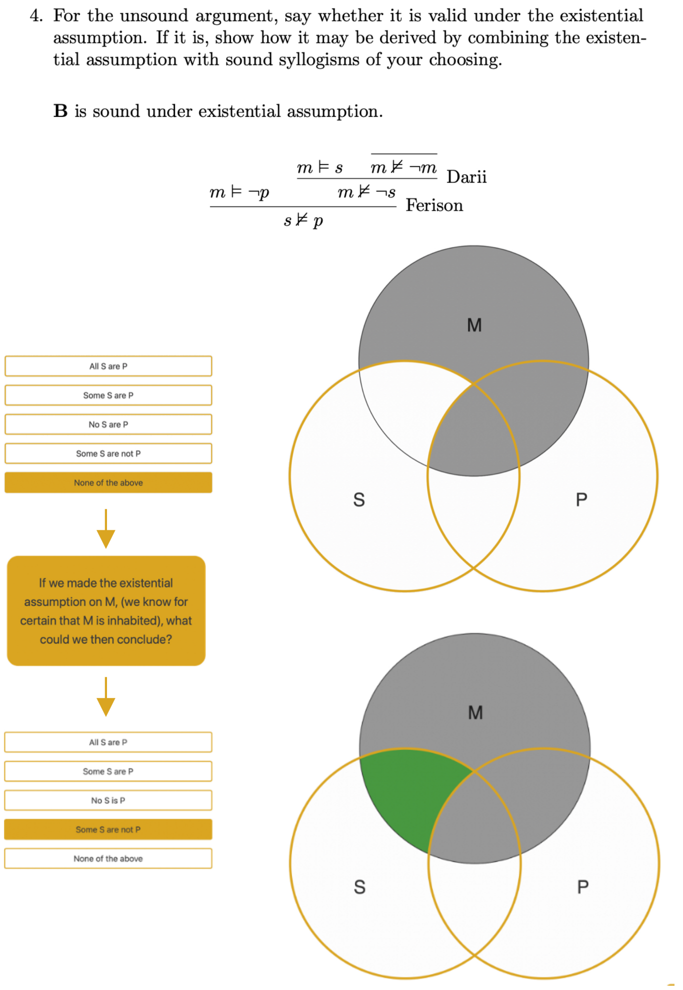

Figure 4.3: How Syllogism Studio could be used to answer an question from Inf1A

The goals left to be addressed are 1 and 2. While the web application was designed to satisfy these goals, we cannot be certain of their achievement without a user study. A user study that focuses on goal 2 is presented in Chapter 5, and a study on goal 1 in Chapter 6.

4.4 Limitations

A limitation of Syllogism studio is that it does not have any natural language processing capabilities, simply substituting the terms entered as S, M, and P without any processing. As explained by Howard [20], this becomes an issue when the predicate of a premise is an attribute of the subject. For example, take the syllogism:

> All feathers are soft
> Some feathers are white
> Therefore, some white things are soft

In the minor premise, "white" is used as an adjective, describing an attribute of "feathers". When "white" appears as the subject of the conclusion, it is naturally converted to a noun. However, as Syllogism Studio cannot identify whether terms are adjectives or nouns, the conclusion offered by the app would be "Some white are soft", as shown in figure 4.4.

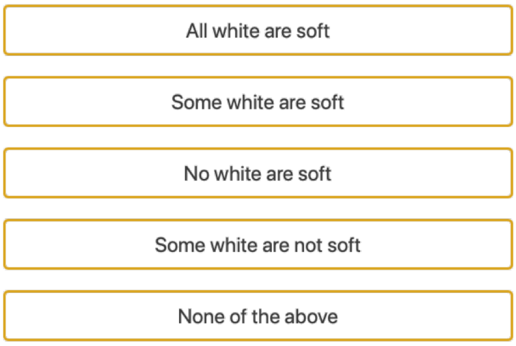

Figure 4.4: Grammatically incorrect conclusions

A similar issue is found when the predicate of a proposition is a verb, such as in figure 4.5. To enter the major premise into the app, the verb "fly" would first have to be translated into an adjective, and the premise entered as "No penguins are flying". If the term "flying" then appeared as the subject of a premise, the app would produce a grammatically incorrect statement. Syllogism Studio's inability to identify parts of speech means it can only handle letters, or plural verbs as terms without making grammatical errors.

B.

No penguins fly.
All penguins are birds.
$\therefore$ Some birds do not fly.

with p = penguins; f = fly; b = birds.

Figure 4.5: Sample question from Inf1A

Another limitation of Syllogism Studio is how it displays syllogisms with two particular premises. When expressing syllogisms in triple Venn diagrams, areas that are known to be inhabited in a double Venn diagram reach across two regions. The green, inhabited areas now denote pairs of regions, at least one of which is inhabited. When two such pairs are represented on a triple Venn diagram, they may intersect, or be adjacent to each-other. It is then difficult to distinguish which regions are paired together to denote a larger, inhabited region, as shown in figure 4.6. For this reason, Syllogism Studio best caters for premise pairs with a valid conclusion, as they do not have two particular premises.

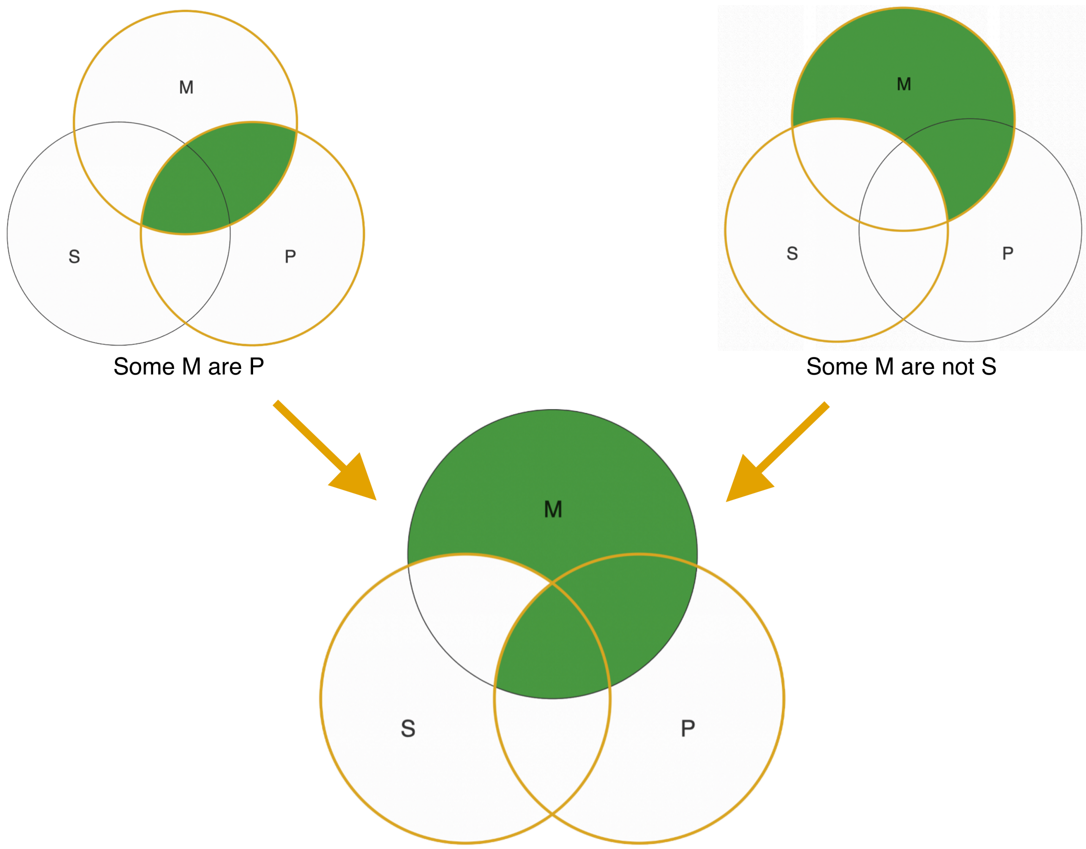

Figure 4.6: Venn diagram displaying two particular premises

4.5 Chapter Summary

This chapter provided an overview of how Syllogism Studio was created, what the design goals were, and what steps were taken to meet them.

What is Syllogism Studio?
Syllogism Studio is a web application that provides an interactive step-by-step guide to using Venn diagrams as a way to represent, and interpret, syllogisms.

Can I use it?
Yes, you can access the tool at `https://syllogism.studio`.

Is the tool perfect?
Unfortunately not. The tool doesn't have any language processing capabilities currently, and does not display overlapping or adjacent inhabited areas well. If you come across any other limitations, you can use the feedback form on the app to raise the issue.

Chapter 5

Evaluation

This chapter documents a study to evaluate Syllogism Studio's usability, and whether its use reduces the influence of non-logical factors when reasoning. The original intention was to conduct a think-aloud usability test, followed by an investigation into how Syllogism Studios affects syllogistic reasoning. However, due to the time lost waiting for ethics approval, the studies were combined. This allowed time for the results to be analysed, and any changes to the web application to be made by the end of the semester.

5.1 Goals

The overall goal of the combined study was to answer the following questions:

1. How does the Venn diagram method presented by Syllogism Studio affect users' syllogistic reasoning?
    (a) Does using Syllogism Studio whilst solving syllogisms increase or decrease the proportion of correct answers?
    (b) Is the atmosphere effect stronger or weaker when using Syllogism Studio?
    (c) Does Syllogism Studio reduce errors due to language ambiguity?
2. Is the application usable?
    (a) Do participants report difficulty in using the application?
    (b) What changes should be made based on user feedback?

5.2 Experimental Design

To investigate how Syllogism Studio affects users' syllogistic reasoning, participants were asked a series of questions on syllogisms to be answered with, or without the use of the web app. The methodology was based on the work of Johnson-Laird and Steedman, in their investigation into the psychology of syllogisms [22].

Two experiments were conducted by Johnson-Laird and Steedman to investigate how the mood of the premises affect the conclusions drawn by participants. In their first experiment, on which my methodology is based, 20 undergraduate students were presented with pairs of premises from which valid conclusions could be drawn. A total of 27[1] premise pairs that had a valid conclusion were included. The premises were presented as statements regarding a group of people that the participants should imagine as assembled in a room. The terms of the premises could then be interpreted as characteristics of these people. Based on the information provided, participants were expected to make deductions, and write down a statement that followed from the premises.

The participants received the premise pairs in the form of a paper booklet, where each question was mimeographed and assembled into a booklet in random order. The responses, and their frequency, were then presented by premise pair. Since their investigation concerned syllogisms that were not in standard form, the findings themselves are not relevant to this study. However, their investigation into bias in the conclusions drawn makes the methodology particularly valuable.

Methodology

Sixteen participants were presented with sixteen premise pairs and asked to derive conclusions. Participants answered eight questions without any additional support, and eight questions whilst using Syllogism Studio. The conclusions were presented as multiple-choice, rather than presenting open-ended questions as was done by Johnson-Laird and Steedman to align with the interface of Syllogism Studio.

In Lefford's study on the emotional content of syllogisms [23], the order of the two sections of the study affected the results. To control for this order bias, group A answered questions with the web application first, then went on to answer without. Group B completed the sections in the opposite order.

To further reduce the risk of order biases, every participant was presented with the questions in a different order. The question order was drawn from a 16x16 Sudoku puzzle, where each row in the puzzle corresponded to a participant, and the order in which the numbers 1 to 16 appeared in that row determined the order in which questions would be presented. It was then possible to ensure that every question would be answered with and without Syllogism Studio exactly eight times.

The questions were printed and assembled into paper booklets along with sample questions and instructions.

Participants

The participants in this study were undergraduate students at The University of Edinburgh, that did not study informatics, and had no prior training in logic. An ideal participant selection criteria would be undergraduate informatics students that had no

[1]When presenting the premises in standard form, only 19 premise pairs have valid conclusions, as shown in Appendix A. Since Johnson-Laird and Steedman did not enforce standard form in their study, the number of premise pairs with valid conclusions rose to 27.

prior training in logic as this is the target audience of Syllogism Studio. However, since all informatics students take Inf1A in their first semester, very few, if any, potential participants would have met these criteria.

The intention was to give participants a brief introduction to syllogisms in small groups before they went on to answer the questions. This would be similar to an introductory class that Inf1A students might be given. However, as Covid restrictions did not allow for in-person meetings, a brief introduction was provided in the form of sample questions instead.

Question Design

Since previous studies have observed participants performing better with sound syllogisms than unsound syllogisms [15] [32], the study focused on sound syllogisms. This was to avoid the risk of participants finding the questions too difficult to answer and simply guessing instead of reasoning to find a conclusion.

Conditionally sound syllogisms were omitted, as an understanding of the existential assumption could not be expected from participants seeing syllogisms for the first time.

This resulted in the 15 unconditionally sound syllogisms being the focus of the study. One unconditionally sound syllogism was omitted, to be used as a sample question on the instruction sheet, along with one example of an unsound syllogism. These two sample questions were selected as they were the easiest to answer, according to the results of Johnson-Laird and Steedman.

This left 14 sound syllogisms. For an even distribution of syllogisms across the four figures, one first figure and one fourth figure unsound syllogism was included, so that there were four questions for each mood.

Since testing for all four non-logical factors discussed was beyond the scope of what was reasonable to study in a single experiment, it was decided to focus on only two factors - the atmosphere effect, and language ambiguity. Belief bias and the emotional content of syllogisms were controlled for by using terms that would not invoke emotion or prior belief. “Scientists”, “Mathematicians” and “Professors” were selected as the terms, as it removed the need for participants to identify S, M, and P themselves. When using the web app, they could identify the S M and P terms by taking the first letter of each term.

Experimental Interface

Syllogism Studio was adapted in the following ways before being presented to participants:

1. No feedback was given on whether participants had selected the correct answer in the app.
2. Discussion of the existential assumption was removed, as it was not relevant to the study

3. The introductory text was adapted to tell participants which terms were S M and P, as understanding of standard form was not being assessed.
4. An anonymous feedback form was added to allow users to report any difficulties they experienced whilst using the app.

5.3 Data Collection

Ethics approval for this study was sought in February. After an initial follow-up email, no response was received from the ethics committee, and therefore ethics approval was not obtained. For this reason, careful consideration was given to not collecting any personal data or identifiable information from the participants. To preserve each participant's anonymity, the following steps were taken:

1. The materials for each participant were placed in identical envelopes. These envelopes were distributed at random, so that I would not know which participant filled in each booklet.
2. The materials were collected back in a sealed envelope.
3. Once the envelopes were all returned, they were opened, and assigned an ID. The ID was linked to the participant's answers, but not to their identity. As the envelopes were all opened at the same time, there was no way to link results to a particular participant.
4. The results were aggregated by premise pair, so that a single participant's responses could not be identified.

5.4 Results

The aggregated responses are presented in two tables. Table 5.1 presents participants' answers whilst using the web app, and table 5.2 without. Each cell in the table contains the conclusions drawn by participants when presented with the premises in the corresponding row and column. All of the responses that were received are presented, with the frequency of each response given in brackets after the statement. The responses within a cell are ordered from most, to least frequent, with the correct response in bold font.

The frequency of responses in each cell should total 8, as 8 participants were asked to answer each question with, and without, the web app. However, one participant failed to return the booklet, meaning that only 7 responses are logged in some cells. One participant failed to give a response to a question, and so the premise pair "Some P are not M" and "All M are S" received only 6 responses.

Table 5.1: Participant responses with the use of Syllogism Studio

	Major Premise							
Minor Premise	All M are P	Some M are P	No M are P	Some M are not P	All P are M	Some P are M	No P are M	Some P are not M
All S are M	**All S are P (6)** Some S are P (1) Some S are not P (1)		**No S are P(8)**				**No S are P(7)** None of the above (1)	
Some S are M			**Some S are not P (7)** None of the above (1)				**Some S are not P (4)** Some S are P (3) None of the above (1)	
No S are M					**No S are P(7)** None of the above (1)			
Some S are not M	Some S are not P (4) Some S are P (2) No S are P (1) **None of the above (1)**				**Some S are not P (6)** Some S are P (1) No S are P (1)			
All M are S		**Some S are P (5)** Some S are not P (1) None of the above (1)		**Some S are not P (4)** Some S are P (2) None of the above (1)		**Some S are P (4)** Some S are not P (2) None of the above (1)		Some S are not P (3) **None of the above (2)** Some S are P (1)
Some M are S	**Some S are P (5)** Some S are not P (1) None of the above (1)		**Some S are not P (4)** Some S are P (2) No S are P (1)				**Some S are not P (4)** Some S are P (3)	
No M are S					**No S are P (7)**			
Some M are not S								

Table 5.2: Participant responses without the use of Syllogism Studio

	Major Premise							
Minor Premise	All M are P	Some M are P	No M are P	Some M are not P	All P are M	Some P are M	No P are M	Some P are not M
All S are M	**All S are P (5)** Some S are P (1) None of the above (1)		**No S are P(6)** Some S are P (1)				**No S are P(7)**	
Some S are M			**Some S are not P (6)** None of the above (1)				**Some S are not P (5)** Some S are P (1) No S are P (1)	
No S are M					**No S are P(7)**			
Some S are not M	Some S are not P (5) Some S are P (1) **None of the above (1)**				**Some S are not P (7)**			
All M are S		**Some S are P (6)** None of the above (2)		**Some S are not P (5)** Some S are P (2) None of the above (1)		**Some S are P (6)** Some S are not P (1) None of the above (1)		Some S are not P (5) Some S are P (2) **None of the above (1)**
Some M are S	**Some S are P (7)** Some S are not P (1)		**Some S are not P (5)** Some S are P (2) None of the above (1)				**Some S are not P (5)** Some S are P (1) No S are P (1) None of the above (1)	
No M are S					**No S are P (5)** Some S are not P (2) None of the above (1)			
Some M are not S								

5.5 Discussion

When comparing the results to those of Johnson-Laird and Steedman [22], the participants performed better than expected.

The valid conclusions were drawn most frequently for all premise pairs, other than the two premise pairs with no valid conclusion. For these two unsound syllogisms, the most frequent responses aligned with the atmosphere of the premise. This is true of the valid conclusions too, meaning that it is difficult to tell if the participants reasoned to reach their valid conclusions, or made a correct guess based on atmosphere.

As well as atmosphere, the multiple-choice format of the question may have contributed to the high proportion of correct responses. Note that in Johnson-Laird and Steedman's study, the participants were expected to draw their conclusions with no assistance. While this has meant that there are fewer errors in participants' reasoning to study, it suggests that presenting multiple choice conclusions, as is done in Syllogism Studio, can help users draw valid conclusions.

1. **How does the Venn diagram method presented by Syllogism Studio affect users' syllogistic reasoning?**

   (a) **Does using Syllogism Studio whilst solving syllogisms increase or decrease the proportion of correct answers?**

	Total responses	Correct responses	Incorrect responses	Accuracy
With Syllogism Studio	119	81	38	68%
Without Syllogism Studio	120	84	36	70%
Overall	239	165	74	69%

Table 5.3: Summary of correct and incorrect responses

   Using Syllogism Studio does not appear to have any effect on the accuracy of participants' reasoning. While the overall accuracy is slightly lower whilst using Syllogism Studio, the difference is negligible and is not statistically significant ($p > 0.1$, left-tailed test).

   (b) **Is the atmosphere effect stronger or weaker when using Syllogism Studio?**

   The two premise pairs with no valid conclusions drew incorrect responses that were consistent with the atmosphere effect. Both pairs had one universal affirmative (A), and one negative particular (O) premise. Under the atmosphere effect, it is expected to see a bias towards selecting a negative particular conclusion.

   Whilst using Syllogism Studio, half of the total responses considered were consistent with the atmosphere effect. Without the assistance of the web app, this rose to two-thirds.

	Total responses	Incorrect responses	Incorrect responses consistent with atmosphere
With Syllogism Studio	14	11	7
Without Syllogism Studio	15	13	10
Sum	29	24	17

Table 5.4: Atmosphere effect in premise pairs with no valid conclusions

This provides some evidence towards Syllogism Studio decreasing the likelihood of drawing an invalid conclusion due to atmosphere. However, as only two premise pairs with a total of 29 responses were considered in this analysis, the sample is too small to generalise the results.

(c) **Does Syllogism Studio reduce errors due to language ambiguity?**

Four premise pairs were identified as cases where language ambiguity led to incorrect conclusions.

One premise pair received responses of the type "Some S are P", when only "All S are P" follows. Here, a reasoner would have incorrectly made the existential assumption on S, likely due to ambiguity in the meaning of "All".

Three premise pairs had incorrect responses due to reasoners inferring "Some A are not B" from "Some A are B", or vice versa.

In both premise pairs with no valid conclusion, some participants converted the particular negative premise to the particular affirmative equivalent, before drawing a conclusion, leading to incorrect responses.

For the premise pair "Some P are M" and "Some M are S", the valid conclusion is "Some S are not P". However, some participants responded "Some S are P".

The responses to these four premise pairs are summarised in Table 5.5.

	Total responses	Incorrect responses	Incorrect responses due to language ambiguity
With Syllogism Studio	29	16	6
Without Syllogism Studio	31	18	6
Sum	60	34	12

Table 5.5: Incorrect responses due to language ambiguity in four premise pairs

Using Syllogism Studio did not affect the number of incorrect responses due to language ambiguity. Even if there was an effect observed, the small sample size may have made it difficult to generalise the results.

2. **Is the application usable?**

   (a) **Do participants report difficulty in using the application?**

   As the study was conducted remotely, without observing participants' screens, usability data was only gathered through the feedback form on the app.

   Three responses were received from the feedback form. Two of the participants that submitted the form reported difficulty with understanding the colours in the Venn diagrams, as intuitively, grey and red both implied that a region was uninhabited.

   Two feedback forms mentioned finding the breakdown of how to work through a question helpful.

   (b) **What changes should be made based on user feedback?**

   In response to the feedback, green was used to display inhabited regions, instead of red. While the figures in this report have used green to denote inhabited areas throughout, red was used up until the user feedback was received.

5.6 Design limitations

Using the web application had no significant effect on participants' reasoning. It was expected that using Syllogism Studio would either increase or decrease the overall accuracy of participants' validity judgements. It seemed reasonable to hypothesise that if the app was successful in helping users interpret syllogisms, that the accuracy would increase. If the application was confusing or misleading, the accuracy would decrease. However, using the web application had no significant effect on participants' reasoning. Does Sylliogisms Studio, somehow, achieve nothing?

The lack of observable effect and the small number of feedback forms submitted suggest an alternative explanation - that the participants did not use the web app to assist with answering the questions.

Follow-up conversations with participants revealed that whilst they did use the app, many scrolled through the Venn diagrams without trying to understand them, then based their answers solely off of the premises on the question booklet, and not the Venn diagram.

This most likely happened as participants would have needed to put significant effort into understanding the Venn diagram method, and most, being able to answer the questions without it, chose not to.

The limitation here is that participants were not explicitly taught the Venn diagram method before the study, and so using an app that displayed Venn diagrams was not beneficial to them. As the study was conducted remotely due to Covid restriction, it was not possible to support the participants in using the web app and interpreting the Venn diagrams.

It would have been better to select participants that were familiar with the Venn diagram method for interpreting syllogisms, such as current Inf1A students. This was not possible given the timing of the study, as Inf1A was taught in semester 1, while the study was conducted in semester 2.

The second limitation is with the question selection. The questions that had the most valuable responses in terms of analysis were those with premise pairs that had no valid conclusion.

It was difficult to know which questions would give insightful responses before presenting them to participants. However, if the study was done again, it would be best to include 2 sound, and 2 unsound syllogisms from each figure, rather than focusing on sound syllogisms. This would likely lead to more erroneous responses, and therefore more data on the errors that users typically make. The sample of invalid conclusions would then be large enough for the result on the atmosphere effect to be generalised, and for some effect on language ambiguity to be observed.

The final limitation of this study is that the feedback form was included at the bottom of the web app. Participants that did not fully engage with the app would not have seen it, and therefore did not fill it in. As a result, limited data was gathered on usability. It would have been better to include short questions on usability as part of the question booklet, to encourage more users to provide feedback.

5.7 Chapter Summary

This chapter documented a study conducted to investigate how Syllogism Studio affects users' syllogistic reasoning.

What did the study entail?
sixteen participants were presented with sixteen questions on syllogisms in a paper booklet. They were asked to answer eight questions with, and eight without, the use of Syllogism Studio.

What were the results?
The participants answered more questions correctly than incorrectly. The proportion of correct and incorrect responses were consistent both with and without the help of Syllogism Studio. There was some evidence that Syllogism Studio reduced the atmosphere effect, but there was no evidence that it reduced the number of errors due to language ambiguity.

Were the any limitations to the study?
Yes. Due to the limitations outlined in Chapter 5.6, participants did not engage with the web app fully. Because of this, the results cannot be taken as a true reflection of Syllogism Studio's effect on users' ability to reason with syllogisms.

What do these limitations mean for the project?
The limitations mean that reliable data on the usability and effectiveness of the app was not gathered. To address this issue, a second user study was conducted, outlined in Chapter 6.

Chapter 6

User Study with Inf1A students

This chapter outlines a second user study, which was conducted to address some of the limitations of the experiment described in Chapter 5.

6.1 Project Presentation

The user study was conducted informally during the honours project poster presentation day, an online event attended by informatics staff and students.

Informatics students that showed interest in the project were invited to visit Syllogism Studio and test the app using sample syllogisms, or by picking their own examples to try. These students were primarily first and second year undergraduates, and so were familiar with syllogisms from taking Inf1A in previous semesters.

As the participants used the app, they were asked to share their thoughts aloud. With multiple users testing the app at the same time, this led to a group discussion on the usability of the app. When only one student was present it allowed a more focused discussion on how they interpreted the Venn diagrams drawn by the app to find a conclusion. Note that the feedback on whether a user had selected a correct or incorrect conclusion was hidden during the presentation, to promote discussion on a conclusion's validity.

The study followed the needs and interests of the users, resulting in a session that was part focus group, part think-aloud study. No personal information or other data was gathered from those that tested the app, other than what was submitted through the feedback form. Completion of this form was entirely optional and anonymous.

6.2 Feedback

After testing the app, participants were asked to fill in an anonymous feedback form. A total of 6 participants submitted feedback in the form of written responses to 3 open-ended questions. The questions and full responses are found in Appendix B. A summary of the key points from the feedback forms and verbal feedback is summarised as follows:

1. The web app was well-received by previous Inf1A students. Many expressed that they would have found the app useful during the course.
2. For those already familiar with syllogisms and the Venn diagram method, the app is easy to understand and use.
3. The app displays well on a range of web browsers but is not as successful on different screen sizes. Particularly wide or narrow windows/screens cause layout issues.
4. Some users like the simple grey/green colouring of regions in the Venn diagram, whilst others prefer to use ticks and crosses to annotate regions.
5. There is a demand for more topics within Inf1A to be added.

The feedback received suggested that Syllogism Studio would be a valuable tool for Inf1A students. While there are no quantitative measures of how well Syllogism Studio would support future Inf1A students, the overall positive response is encouraging, and suggest that the app would cater well for Inf1A students' needs.

The feedback also highlighted usability issues that, along with the design limitations outlined in Chapter 4.4, should be used to continue to improve the application.

Chapter 7

Conclusion

This report discusses three main bodies of work: background research, design and implementation, and evaluation.

The background research summarised the relevant literature on syllogistic reasoning, with a focus on how and why reasoners make errors. It was argued that using Venn diagrams to assist with reasoning could reduce the frequency of errors. As a result, a web application was designed to support the use of Venn diagrams as reasoning tools for syllogisms. The application can be accessed at `https://syllogism.studio`. The application is freely available to all, although it is designed specifically for informatics students studying syllogisms as part of Inf1A.

Syllogism Studio was then used to evaluate the hypothesis that using Venn diagram representations of syllogisms could reduce common errors in syllogistic reasoning. The study provided little evidence either for or against the hypothesis. However, the lessons learned from the study could be used to conduct a more informative investigation in the future.

Finally, the application was presented to informatics students that had recently taken Inf1A. This final element of evaluation provided feedback on whether the application succeeded in catering to the needs of informatics students, and provided ideas on how the application could be extended.

There are two main aspects of this project that invite further work: the development of Syllogism Studio, and the investigation of its effect on syllogistic reasoning.

Syllogism Studio could be further developed by adding language processing capabilities, developing a notation for intersecting or adjacent inhabited regions, and ensuring usability across a range of screen sizes.

The experiment outlined in Chapter 5 should be repeated with changes to the questions and participant selection criteria. More unsound syllogisms should be incorporated, to be presented to participants that are currently studying Inf1A. Based on the results of this modified study, an attempt should be made to prove or disprove the hypothesis that informed the design of Sylloigism Studio - that the use of Venn diagrams can reduce the errors made due to non-logical factors in syllogistic reasoning.

Bibliography

[1] The pudding. Online. Available at https://pudding.cool.

[2] Reasoning - syllogisms. Online. Available at https://www.tutorialspoint.com/reasoning/reasoning_syllogism.htm.

[3] Syllogisms test. Online. Available at https://www.fibonicci.com/logical-reasoning/syllogisms-test/.

[4] Valid or invalid. Online. Available at https://www.philosophyexperiments.com/validorinvalid/.

[5] Venn diagram generator. Online. Available at https://betterinformatics.com/resources/inf1-cl/venn.

[6] Venn diagrams. Online. Available at https://www.inf.ed.ac.uk/teaching/courses/inf1/cl/tools/venn/.

[7] Raees Aamir. Soak up syllogisms. Online, 2020. Available at http://soakupsyllogisms.azurewebsites.net.

[8] Raees Aamir. Soak up syllogisms: A tool for learning categorical syllogisms. Master's thesis, School of Informatics, University of Edinburgh, 2020.

[9] Aristotle. *The complete works of Aristotle : the revised Oxford translation*. Princeton University Press, Princeton, N.J, 1984.

[10] Bruno G. Bara, Monica Bucciarelli, and Philip N. Johnson-Laird. Development of syllogistic reasoning. *The American Journal of Psychology*, 108(2):157, 1995.

[11] Mike Bostock. D3.js - data-driven documents, 2012. Available at https://d3js.org.

[12] Peter C.-H. Cheng. Graphical notations for syllogisms: How alternative representations impact the accessibility of concepts. *Journal of Visual Languages & Computing*, 25(3):170–185, jun 2014.

[13] Morris Raphael Cohen. *An introduction to logic and scientific method*. Routledge Kegan Paul, London, 1934.

[14] Irving Copi and Carl Cohen. *Introduction to logic*. Macmillan Collier Macmillan, New York London, 1990.

[15] Louis S. Dickstein. Effects of instructions and premise order on errors in syllogistic reasoning. *Journal of Experimental Psychology: Human Learning & Memory*, 1(4):376–384, 1975.

[16] Jonathan Evans. *Human reasoning : the psychology of deduction*. Lawrence Erlbaum Associates, Hove, 1993.

[17] Jonathan St. B.T. Evans, Simon J. Handley, and Catherine N. J. Harper. Necessity, possibility and belief: A study of syllogistic reasoning. *The Quarterly Journal of Experimental Psychology Section A*, 54(3):935–958, aug 2001.

[18] Russel Goldberg. Scrollama.js. Online. Available at `https://github.com/russellgoldenberg/scrollama`.

[19] Russell Goldenberg. Introducing scrollama. Online, November 2017. Available at `https://pudding.cool/process/introducing-scrollama`.

[20] Delton Thomas Howard. *Analytical syllogistics : a pragmatic interpretation of the Aristotelian logic*. Northwestern University studies. Humanities series ; no. 15. Evanston, 1946.

[21] Philip N Johnson-Laird. *Mental models : towards a cognitive science of language, inference, and consciousness*. Harvard University Press, Cambridge, Mass, 1983.

[22] Philip N Johnson-Laird and Mark Steedman. The psychology of syllogisms. *Cognitive Psychology*, 10(1):64–99, jan 1978.

[23] Arthur Lefford. The influence of emotional subject matter on logical reasoning. *Journal of General Psychology*, 34:127–151, 1946.

[24] Yannick Lemieux. Syllog. Online, 2012. Available at `https://syllog.emergence.dk`.

[25] Joshua Little. Syllogism validity checker. Online. Available at `https://mjtlittle.github.io/Syllogism-Checker/`.

[26] Roberta E Love. *Atmosphere effects revisited: The effect of atmosphere on judgments of validity using thematic syllogisms*. PhD thesis, 1998.

[27] Anthony Manser. *Bradley's logic*. B. Blackwell, Oxford Oxfordshire, 1983.

[28] Peter Øhrstrøm, Ulrik Sandborg-Petersen, Steinar Thorvaldsen, and Thomas Ploug. *Classical Syllogisms in Logic Teaching*, pages 31–43. Springer Berlin Heidelberg, 2013.

[29] Peter Øhrstrøm, Ulrik Sandborg-Petersen, Steinar Thorvaldsen, and Thomas Ploug. *Teaching Logic through Web-Based and Gamified Quizzing of Formal Arguments*, pages 410–423. Springer Berlin Heidelberg, 2013.

[30] Peter Øhrstrøm, Ulrik Sandborg-Petersen, Steinar Thorvaldsen, and Thomas Ploug. *Teaching Syllogistics Using E-learning Tools*, pages 103–110. Springer International Publishing, nov 2016.

[31] Helena Patricia Osana. Visualization in syllogistic reasoning. 1989.

[32] James J. Roberge. A reexamination of the interpretations of errors in formal syllogistic reasoning. *Psychonomic Science*, 19(6):331–333, dec 1970.

[33] N. E. Wetherick and K. J. Gilhooly. 'atmosphere', matching, and logic in syllogistic reasoning. *Current Psychology*, 14(3):169–178, sep 1995.

[34] R. S. Woodworth and S. B. Sells. An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology*, 18(4):451–460, 1935.

[35] Robert S. Woodworth. *Psychology*. Methuen, 1945.

Appendix A

Table of Sound Syllogisms

The entries in the table denote the conclusion that follows from the premises in the corresponding rows and columns. Unconditionally valid conclusions are presented in bold text. Conditionally valid conclusion are provided along with their validity condition.

The column for "Some P are not M", and row for "Some M are not S" have been omitted, as there are no valid conclusions with either premise.

	Major Premise						
Minor Premise	All M are P	Some M are P	No M are P	Some M are not P	All P are M	Some P are M	No P are M
All S are M	**All S are P** Some S are P if S exists		**No S are P** Some S are not P if S exists				**No S are P** Some S are not P if S exists
Some S are M	**Some S are P**		**Some S are not P**				**Some S are not P**
No S are M					**No S are P** Some S are not P if S exists		
Some S are not M					**Some S are not P**		
All M are S	Some S are P if M exists	**Some S are P**	Some S are not P if M exists	**Some S are not P**	Some S are P if P exists	**Some S are P**	Some S are not P if M exists
Some M are S	**Some S are P**		**Some S are not P**				**Some S are not P**
No M are S					**No S are P** Some S are not P if S exists		

Appendix B

Responses to Feedback Form

All respondents gave consent for their feedback to be included in this report.

What did you find useful about this web application?

- very clear and informative design
- Really user friendly!
- The way everything is visualised
- Being able to visualise using the Venn diagrams was good
- explained really well how to do syllogisms
- The colours are really useful to help with visualisation :)) I like the grey and green!

Did you find anything difficult to understand?

- not at all!
- "Some S are not P" and "some S are P" can both be true at the same time, right?
- the css messed up all the text boxes and drop downs
- Not really!
- no it was all good really good project
- I think its pretty straightforward

Do you have any suggestions for changes that could be made, or new features that you'd like to see?

- more CL content (Karnaugh maps etc)
- that white[1] screen thingy
- add existential assumptions

[1] This appears to be a typo, with the text intended to read "wide". Issues with wider screen sizes were discussed during the presentation

- often forget whether the coloured means inhabited or not - a tick or a cross might be good
- More of the CL course implemented in it, explanation of SMP would be good for those who have not done Inf1A and/or CL
- This was kind of discussed at gather town but maybe also having a way which we could also use for our homeworks and questions on paper without having to change colours