Cognitive Defeasible Reasoning

Clayton Baker
University of Cape Town
bkrcla003@myuct.ac.za

Claire Denny
University of Cape Town
dnncla004@myuct.ac.za

Paul Freund
University of Cape Town
frnpau013@myuct.ac.za

CCS CONCEPTS

• **Theory of computation → Logic**; • **Computing methodologies → Nonmonotonic, default reasoning and belief revision**; *Reasoning about belief and knowledge*;

KEYWORDS

propositional logic, defeasible reasoning, cognitive defeasible reasoning, belief revision, belief update, non-monotonicity

1 PROJECT DESCRIPTION

Reasoning is an integral part of human lives. It is well-documented that human reasoning fails to conform to the prescriptions of classical or propositional logic [16], modelling a flexibility considered key to intelligence [14]. The artificial intelligence community therefore seeks to incorporate such flexibility in their work [14]. Non-classical or non-monotonic logic is flexible by nature. Classical reasoning is sufficient to describe systems with a calculated output in an efficient way. However, the way in which humans reason is non-classical, as humans are known to reason in different ways [16]. The problem is that nonmonotonic reasoning schemes have been developed for and tested on computers, but there exists a similarity with human reasoning that needs to be investigated. Our problem is important because we can gain insight into how humans reason and incorporate this into building natural computer systems. An issue which needs to be considered is that humans are diverse subjects - some reason normatively while others reason descriptively. A comprehensive investigation needs to be done to detect which aspects of human reasoning are situated in nonmonotonic reasoning.

2 PROBLEM STATEMENT

2.1 Aims

Various forms of nonmonotonic reasoning have been developed by the AI community, but these same formalisms also have the potential to describe human reasoning. Humans are known to reason nonmonotonically. We want to examine the closeness between how humans and artificially-intelligent computers reason. In particular, we have identified three nonmonotonic reasoning patterns - defeasible reasoning, belief revision and belief update - that we will consider. The direct link between human reasoning and these forms of defeasible nonmonotonic reasoning has not been investigated extensively - our research aims to close this gap. The goal of this project is to test whether, and under what circumstances, defeasible reasoning systems conform to the way in which humans reason. The project involves a theoretical component, in which different forms of defeasible reasoning are considered, and an experimental component, in which human subjects are provided with case studies of reasoning and their responses are measured against various forms of defeasible reasoning.

2.2 Research questions

1. To what extent do the formalisations of Defeasible Reasoning, Belief Revision and Belief Update match the way in which humans reason?
2. Which properties of Defeasible Reasoning, Belief Revision and Belief Update do not hold with human reasoning?
3. Which of Defeasible Reasoning, Belief Revision and Belief Update is a closer representation of human reasoning?

These three forms of non-classical reasoning have been described formally in terms of properties that have been deemed desirable for them to model. We thus propose to divide this question into three parts, with each of us investigating the correspondence between the formal and the cognitive for one of the reasoning systems in question.

2.3 Requirements

This project entails research rather than Software Engineering. The project will be divided into three parts, corresponding to the three forms of nonmonotonic reasoning that we will consider. A critical component of our project is the conducting of a suitable survey with a random sample of human subjects who speak and understand English. We will present our subjects with three surveys - one each for defeasible reasoning, belief revision and belief update - which contain examples by which the reasoning patterns of our subjects are tested. Further requirements and details will be discussed in section 3.

3 PROCEDURES AND METHODS

3.1 Experiment design, implementation strategy and expected challenges

We will not be undertaking a Software Engineering project and thus we will not develop an end system or prototype. Instead, we have an experiment which we will conduct with human subjects (participants). Our experiment involves two phases: the design and the execution of a survey for each of the three reasoning patterns. Each survey will contain a section in which the participants will fill in their personal contact details. Each participant's survey will be treated with confidentiality and integrity. The results of the survey will be communicated to each participant and we will also provide our contact details to each participant in the event that they have any questions about our research.
The survey will also include different types of examples relating to the three reasoning patterns we are investigating. The types of questions in the survey will include real-world concrete examples, in which the participant will answer with 'Yes' or 'No', and also provide a reason for their answer. An example of a real-world question would be: "If Lee-Anne has a cake to bake, will she use an oven?". Other questions will contain more abstract examples which test

properties of the reasoning pattern in question. An example of an abstract question would be: given the following, "If A then B", and "If C then A", can we say that "If C then B?". The benefit of abstract examples is discussed in [14].

The reason for the answer given by the respondent is important in our experiment. We will use it to detect the reasoning style that the participant used and categorise it as normative or descriptive. A normative reasoning style makes a claim about how something should be whereas a descriptive reasoning style makes a subjective, detail-oriented observation about a particular situation. The 'Yes'/'No' answers are of quantitative importance and can be evaluated using statistical methods - such as the variation in terms of mode across the three reasoning forms - on our population sample. Some quantitative outcomes that we are interested in are the proportion of respondents that reason normatively and the proportion of respondents that reason descriptively.

We have chosen a population sample of 30 respondents for each of the three surveys. Having investigated literature detailing similarly designed experimental studies, we note that this is a relatively large sample size. This number has been chosen because this project has a time-span of approximately 5 months. In this period, we wish to design a survey in which the respondents will be able provide high-quality answers so that the results we obtain will still be meaningful. In the event that some results do become spoiled, there would be enough remaining for us to use.

We have considered three platforms on which to conduct our survey. The first is Mechanical Turk [13]. It is an online website, hosted by Amazon, which allows registered users to answer surveys virtually for a monetary reward. A second option is to hold a focus group in a controlled venue, such as a lecture theatre in a university. Three focus group sessions would be needed in order for all three surveys to be conducted. Each participant in the focus group would be compensated. A third option would be to publish our survey on Google Forms and distribute it via channels such as the undergraduate computer science mailing list at the University of Cape Town. Respondents would be given the choice to complete the survey and receive compensation, but co-operation is not guaranteed.

We prefer to conduct our survey on Mechanical Turk. This method provides a wide range of respondents - all registered users of Mechanical Turk. Another reason to prefer Mechanical Turk is that it provides a more representative sample of respondents as opposed to only university students. This platform also gives us control over who we select as respondents before the survey is distributed: on Mechanical Turk we have the option to choose, for example, the age group, gender and education level of our respondents. The platform is familiar to the global research community and eliminates the need for arranging a venue and distributing physical surveys, which we would need to do if we were to conduct our survey in a focus group. Google Forms would be our second choice of platform because it does not have the logistical complexity that a focus group has. Participants such as university students, however, may not be interested in our survey and incentives cannot be directly implemented using Google Forms. Should both Mechanical Turk and Google Forms not yield our desired responses, we shall canvass students at the University of Cape Town to participate in a focus group. More planning will be required in this case.

3.2 Testing and evaluation

Our project is two-fold, comprising both the development of suitable questions and the conducting of three surveys. The questions developed will be tested by means of a trial set of surveys conducted using a small sample of people as respondents. This is, in part, to test the process involved in using Amazon's Mechanical Turk before entrusting the whole experiment to be done by means of it, but will also give us insight into how our questions may be received and interpreted. Additionally, we will ask a mix of experts and laypeople, selected by convenience, to evaluate the proposed surveys in terms of coherence, clarity and other desirable characteristics of questions, more examples of which are found in [11]. This will be done in a survey style and will have a space for additional comments to encourage further commentary from these participants. Participants in the final set of surveys will be asked to give overall feedback on the experience. This is included to question, and facilitate their reflection on, the authenticity of their responses, given that what people say they would do does not always match up with what they do in reality. Suggestions of things on which to comment will be provided - for example, how concrete they considered our examples to be.

Exploratory data analysis such as correspondence analysis will be performed on the final data set. In the case of correspondence analysis, this will be to evaluate if the associations between questions and their answers match overall with what we were expecting based on the formalised reasoning patterns.

Responses to survey questions will comprise two parts: a closed aspect, for example a binary response, and a leading aspect, in this case an explanation of the response to the first part. For the questions that have a binary response (yes/no questions) aspect, qualitative measures, such as the mode and hit rate, will be used for evaluating its success. In this project, we propose to qualify a hit as an instance of the response expected by the relevant formalism matching the response observed. Qualitative analysis will be applied to the leading aspect. This entails a multi-step process involving the eight-step process of coding the data, as outlined in [2].

Ultimately, we will be interpreting the results and evaluations for the purpose of deciding whether there is evidence to support the formalisms having a normative or a descriptive relationship with the reasoning patterns of humans.

3.3 Theoretical contributions

The formal systems against which we shall be evaluating their conformance to human reasoning are all extensions of propositional logic. Each of them introduces certain non-truth functional extensions to propositional logic by introducing new operators, and, in the case of defeasible logics of the KLM [10] approach, a new notion of logical consequence is also introduced. In all approaches, the behaviour of the extensions is specified by a set of axioms.

For each example used in the survey, there is thus a two-fold process of confirmation that must occur in order for its relevance to the project to be demonstrated. The first is a modelling process - the examples must be formalised in the language of the extended propositional logic. Here we shall be guided by the usual translations between English and the logics in question. For example, the formal equivalent of "α if γ" is usually accepted as $\gamma \rightarrow \alpha$ and in the defeasible logic community the informal translation of $\alpha \mid\!\sim \gamma$ is usually taken to be something along the lines of "if α then typically γ."

Once the example has been modelled in the formal language, the next step is to derive from the axioms governing the extensions to the logic what the 'correct' answer is for the given example; i.e. what the prediction of the approach in question would be for the example. Each example is thus to be accompanied by two pieces of theoretical confirmation: first a formal model of the example, and second a proof deriving the prediction of the approach in question for that example.

4 ETHICAL, PROFESSIONAL AND LEGAL ISSUES

Ethical issues have been identified in the experimental component of the project, as it requires responses to surveys from human subjects. Prior to the running of the experiments, we shall therefore need to obtain ethical clearance from the Faculty of Science Human Research Ethics Committee. This shall be sought by submitting an ethics application form to the committee before the experiments take place.

When the survey is conducted, respondents will first be presented with instructions, a description of the project and the role they would play in that, their rights as participants, and a digitised consent form in the style of that which the University of Cape Town's ethics application includes. Contact details of the researchers will also be provided. Before data-handling, all survey responses shall be anonymised, excluding factors specified to indicate the respondent's membership of our experiment's desired population.

There are no apparent professional issues as this is purely a research project, and as such we have no 'client' in particular.

The intellectual property for the final reports shall reside in the authors of the reports. These reports will be made available to the survey respondents that grant us permission to contact them. Ownership of the questions used in the survey will be provided to the University of Cape Town.

5 RELATED WORK

Humans are known to reason about situations in everyday life. Nonmonotonic reasoning is the study of those ways of inferring additional information from given information that do not satisfy the monotonicity property which is satisfied by all methods based on classical logic [10]. With nonmonotonic reasoning, a conclusion drawn about a particular situation does not always hold. This type of reasoning is described in the context of AI [15].

5.1 Defeasible Reasoning

In philosophy, when a conclusion has the potential to be withdrawn, or when a conclusion can be reinforced with additional information, the conclusion is said to be defeasible.Defeasible reasoning occurs when the evidence available to the reasoner does not guarantee the truth of the conclusion being drawn [14]. A defeasible statement has two identifiable parts: an antecedent and a consequence [6]. In the statement "If A then B", "A" would be the antecedent and "B" would be the consequence. Often, the meaning attached to a consequence, given an antecedent, is not straightforward. We shall now illustrate this with an example. Consider the following statements: "employees pay tax" and "Alice is an employee". From the statements given, can we conclude that "Alice pays tax"? Using defeasible reasoning, we can infer that "Alice pays tax" and "Alice does not pay tax". The conclusion "Alice pays tax" depends on whether Alice is a typical employee or whether Alice is an exceptional employee and thus does not pay tax.

5.2 Belief Revision

Belief revision is a form of belief change [9]. Belief change involves a belief base and a belief set [4], where explicit knowledge the agent has about the world resides in the base and inferences or knowledge derived from that in the base resides in the belief set. In belief revision, conflicting information indicates flawed prior knowledge on the part of the agent, forcing the retraction of conclusions drawn from it [9, 12]. Information is then taken into account by selecting the models of the new information closest to the models of the base, where a model of information μ is a state of the world in which μ is true [9].

An example of this reasoning pattern will now be described. Consider the same statements used above in the defeasible reasoning example. Using the reasoning pattern of belief revision, we can infer from our beliefs that Alice does pay tax. Suppose that we receive the additional information that Alice does not pay tax. This is inconsistent with our belief base, so a decision must be made regarding which beliefs to retract prior to adding the new information into our beliefs. We could revise our beliefs to be that "employees pay tax" and "Alice does not pay tax". In [3], this decision is proposed to be influenced by whether we believe some statements more strongly than others. In [1], it is proposed to be influenced by closeness or the concept of minimal change - we are simply aiming to change as little about our existing knowledge as we can do without having conflicting beliefs.

5.3 Belief Update

Belief update is another form of belief change. In it, instead of conflicting information indicating a mistake on the part of the agent, rather the conflicting information is seen as reflecting the fact that the world has changed (without the agent being wrong about the past state of the world).

To get an intuitive grasp of the distinction between belief update and revision, take the following example adapted from [9]. Let b be the proposition that the book is on the table, and m be the proposition that the magazine is on the table. Say that our belief set includes $(b \land \neg m) \lor (\neg b \land m)$, that is the book is on the table or the magazine is on the table, but not both. We send a student in to report on the state of the book. She comes back and tells us that the book is on the table, that is b. Under the postulates of belief revision proposed in [1], we would be warranted in concluding that $b \land \neg m$, that is, the book is on the table and the magazine is not. And this seems correct.

But consider now that instead of asking the student to report on the state of the book, we had instead asked her to ensure that the book was on the table. After reporting back that she had indeed ensured that the book is on the table, we again are faced with the new knowledge that *b*. This time adding the new knowledge corresponds to the case of belief update. And here it seems presumptuous to conclude that the magazine is not on the table [9]. Either the book was already on the table and the magazine was not, in which case the student would have done nothing and left, or the magazine was on the table and the book not, in which case the student presumably would have simply put the book on the table and left the magazine similarly so. Crucially these examples are formally identical, so there need be different formalisms to accommodate both cases. In [9], such a formalism is proposed to accommodate the case of belief update.

6 ANTICIPATED OUTCOMES

Major results, including:

6.1 Survey

As part of our project, we will be creating three surveys to establish the correspondence between human reasoning and formal models of nonmonotonic reasoning. The surveys will be the main output of our project. Each of the three surveys will contain both a description section and a consent section. The former will inform the participants about our research. The latter is a form and will be placed at the beginning of the survey, before the survey questions are asked. This is done to ensure our participants are informed and willing to participate in our survey. We will also produce a report for each of the three reasoning patterns. The reports will discuss the findings of the respective survey and suggest the degree to which the results of human reasoning compare to that particular pattern. We anticipate that designing appropriate questions will not be straightforward. In particular, each reasoning pattern has specific properties. For defeasible reasoning, there are six properties [10]. In addition, there are eight properties for both belief revision [8] and belief update [9]. For each reasoning pattern, each property has to be tested in the form of an example. The questions pertaining to these properties also have to be clear enough for our participants to understand. Another challenge that we anticipate is that we do not obtain enough participants for each of the three surveys. We expect that our online survey distribution platform, Mechanical Turk, will provide us with access to at least 30 participants for each survey. We are relying on the Mechanical Turk server and network infrastructure to be fully operational during our project time-span. We also rely on the global community of computer science researchers to be interested and willing to participate in our survey.

6.2 Expected impact

In the defeasible reasoning community, there is currently an emerging research project to test whether the normative properties of formal systems of defeasible reasoning are appropriate for modelling human reasoning, as has been communicated to us by our supervisor. The major impact of the work done in the project will thus be to contribute towards this nascent research paradigm. As far as we are aware, the assumptions underpinning the formalisms of belief revision and belief update have not been empirically tested. Another impact of the work will thus be a first test as to whether the postulated axioms for these domains hold.
Additionally, there are close formal parallels between the postulates for belief revision and defeasible reasoning such that that they may be seen as "basically the same process, albeit used for two different purposes" [5]. While these parallels may hold *formally*, however, they may not hold *empirically*. On a meta-level (i.e. between the individual sub-projects), the projects thus provide a first step to testing whether the translations between defeasible reasoning and belief revision hold empirically. For example, if all postulates for belief revision are found to hold empirically, but some of the postulates for defeasible reasoning are not, given that these postulates are held to be equivalent when viewed in a certain way [7], the results of the projects would suggest that the formal link does not hold empirically.
Finally, the results of the project will be another step towards the greater goal of understanding human cognition.

6.3 Key success factors

This project is a research project, meaning it aims to answer research questions (listed in Section 2.2). The project comprises a theoretical and an experimental component. The theoretical component involving considering various forms of non-classical reasoning; it informs the experimental component. The experimental component has as a key deliverable surveys and responses.

Success factors for the surveys:

- Respondents easily grasp what is required, which is to say that instructions are coherent and clear.
- Respondents feel that the examples are reflective of real life and are not contrived.
- Respondents feel that questions have sufficient answer options.

Success factors for the project as a whole:

- Project plan, and the deadlines therein, adhered to and met.
- Have three surveys, designed adequately to test the three reasoning patterns we are considering.
- Research questions have empirical answers.

7 PROJECT PLAN

7.1 Risks and Risk Management Strategies

See Table 1 in the Appendix.
For each risk identified in the table, there is an impact score given, an assigned probability and both mitigation and contingency plans.

7.2 Timeline

In the appendix, under Figure 1, we show the Gantt chart for our project. Our project begins with a literature review, on the 29th of April 2019, and ends with an open afternoon/evening, on the 15th of October 2019.

7.3 Resources required

The following resources are required by the project team:

- Laptop or other computing device for each person.
- Software required for report writing and statistical analysis.
- Funding to provide monetary incentive for surveys to be taken on Mechanical Turk

7.4 Deliverables and Milestones

During the project life-cycle, there will be evolutionary deliverables. The final deliverables, however, are:

- Survey questions
- Survey results
- A demonstration event
- A final paper, reporting on the project, process and outcomes
- A poster, summarising and reporting on the project
- A website, providing information and documentation
- A reflection paper

Our project milestones begin with the finalisation of our revised project proposal and conclude with a reflection paper. The milestones of our project are listed in full in Table 2 of the appendix.

7.5 Work allocation

The forms of non-classical reasoning our project considers have been allocated among us. The allocation is as follows:
1. Clayton Baker - Defeasible Reasoning
2. Claire Denny - Belief Revision
3. Paul Freund - Belief Update
Each of us will design and conduct our own surveys, according to our allocated reasoning pattern. By comparing the results of our surveys, along with the different theoretical backgrounds that we have, each of us will be able to contribute to answering the research questions given in Section 2.2.

ACKNOWLEDGMENTS

We would like to thank Professor Thomas Meyer for his helpful feedback and guidance with the project topic material.

REFERENCES

[1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. 1985. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic* 50 (1985), 510–530. https://doi.org/10.2307/2274239
[2] John W. Creswell. 2014. *Research design: qualitative, quantitative, and mixed methods approaches* (4 ed.). SAGE Publications, Thousand Oaks, California.
[3] Adnan Darwiche and Judea Pearl. 1997. On the Logic of Iterated Belief Revision. *Artificial Intelligence* 89 (01 1997), 1–29. https://doi.org/10.1016/S0004-3702(96)00038-0
[4] Peter Gärdenfors. 1992. *Belief Revision: An Introduction.* Cambridge University Press, 1–26. https://doi.org/10.1017/CBO9780511526664.001
[5] P. Gärdenfors and D. Makinson. 1994. Nonmonotonic inference based on expectations. *Artificial Intelligence* 65, 2 (1994), 197–245.
[6] Governatori Guido and Paolo Terenziani. 2007. Temporal extensions to defeasible logic. *Australasian Joint Conference of Artificial Intelligence* (2007), 476–485.
[7] S.O. Hansson. 1999. *A Textbook of Belief Dynamics: Theory Change and Database Updating.* Kluwer, Berlin.
[8] Hirofumi Katsuno and Alberto O. Mendelzon. 1991. Propositional knowledge base revision and minimal change. *Artificial Intelligence* (12 1991), 263–294.
[9] Hirofumi Katsuno and Alberto O. Mendelzon. 2003. On the Difference between Updating a Knowledge Base and Revising it. *Belief Revision* 29 (2003), 183.
[10] S. Kraus, D. Lehmann, and M. Magidor. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44 (1990), 167–207.
[11] Jon Krosnick and Stanley Presser. 2009. Question and Questionnaire Design. *Handbook of Survey Research* (03 2009).
[12] João Martins and Stuart Shapiro. 1988. A Model for Belief Revision. *Artif. Intell.* 35 (01 1988), 25–79. https://doi.org/10.1016/0004-3702(88)90031-8
[13] Gabriele Paolacci and Jesse Chandler. 2014. Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science* 25 (2014), 184–188.
[14] Francis Pelletier and Renee Elio. 2005. The Case for Psychologism in Default and Inheritance Reasoning. *Synthese* 146 (08 2005), 7–35. https://doi.org/10.1007/s11229-005-9063-z
[15] John Pollock. 1991. A Theory of Defeasible Reasoning. *International Journal of Intelligent Systems* 6 (1991), 33–54.
[16] M. Ragni, C. Eichhorn, and G. Kern-Isberner. 2016. Simulating human inferences in light of new information: A formal analysis. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI 16)*, S. Kambhampati (Ed.). IJCAI Press, 2604–2610.

APPENDIX

Figure 1: Gantt Chart

Project: CDR			
START DATE	END DATE	DESCRIPTION	DURATION (days)
2019/04/29	**2019/10/15**	**CDR**	**166**
2019/04/29	2019/06/10	*Preparation Phase*	41
2019/04/29	2019/05/02	Submit literature review	3
2019/05/14	2019/05/15	Submit draft project proposal	1
2019/05/16	2019/05/20	Revise project proposal	4
2019/05/21	2019/05/22	Submit final project proposal	1
2019/05/27	2019/05/29	Proposal presentation	2
2019/05/30	2019/05/31	Submit draft ethics application form	1
2019/06/01	2019/06/03	Review of staff feedback after proposal presentation	2
2019/06/04	2019/06/10	Revised proposal finalised and uploaded to Vula	6
2019/06/04	2019/06/10	Follow up on ethics application form	6
2019/06/10	2019/07/05	*UCT Winter Vacation*	25
2019/07/08	2019/08/23	*Project Block*	45
2019/07/15	2019/07/19	Survey feasibility demonstration	4
2019/07/15	2019/07/19	Write Background/Theory section	4
2019/07/15	2019/07/19	Conduct first survey and write-up	4
2019/07/20	2019/07/29	Conduct final survey and write-up	9
2019/07/30	2019/08/02	Write Implementation/Testing section	2
2019/08/03	2019/08/09	Write outline of complete paper	6
2019/08/10	2019/08/16	Write final complete draft of paper	6
2019/08/24	2019/10/15	*Final Phase*	51
2019/08/25	2019/08/26	Submit final project paper	1
2019/08/27	2019/09/02	Submit final survey report	5
2019/09/02	2019/09/16	Final project demonstration	14
2019/09/17	2019/09/23	Create project poster	6
2019/09/24	2019/09/30	Create project webpage	6
2019/10/01	2019/10/07	Write reflection paper	6
2019/10/14	2019/10/15	Open afternoon/evening	1

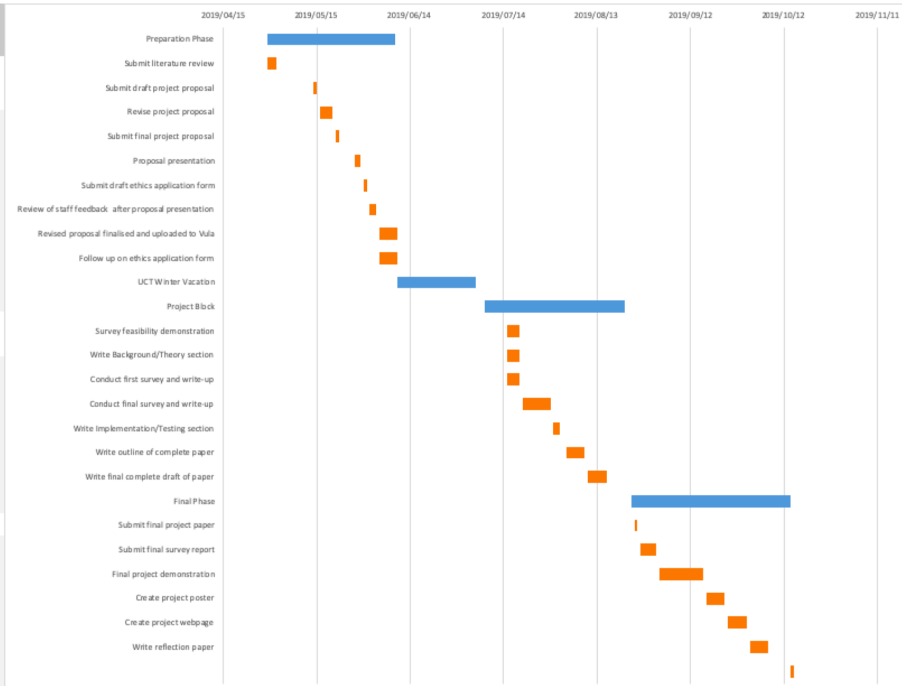

Table 1: Risk Assessment Matrix

Risk	Impact	Probability	Mitigation	Contingency
One or more project member drops out of the year.	Medium	Low	Must be dealt with individually by each project member.	Clear delineation of project responsibilities - each project member essentially runs the whole project by themselves for their domain (one of belief update, revision and defeasible reasoning).
Too few people fill out surveys for relevant results to be obtained (we propose less than 70% of the proposed 30 people as a relevant benchmark).	High	Low	Using Mechanical Turk, which has a reliable way of sourcing survey takers. Appropriate pricing for taking the survey must be researched so as to properly incentive respondents.	In the case that there are not enough people that take the survey on Mechanical Turk, the backup plan is to use Google Forms and for us to individually source people to take the surveys.
Underestimating time needed to complete deliverables, leading to sub-standard work.	Medium	Medium	Proper planning and deadlines set for each deliverable. Individual project members are to be disciplined in ensuring these deadlines are met.	Allocate *more* than enough time for each deliverable.
Experiment design is inadequate for answering research questions.	High	Low	Take time to properly design survey. Before running experiments, ask experts such as supervisor to evaluate whether the proposed survey would be sufficient to meet research goals - could also ask such experts to evaluate after a small trial run (mini-experiment) before final survey. Conduct a trial run on a small number of participants before final survey and anal	None
'Scope creep' - in this case the concern would be that the proposed survey becomes too much work to feasibly implement, be this either due to the amount of theoretical work required for the survey, or that data handling and analysis post-experiment is too much.	Medium	Low	Set concrete goals for work required for survey design and data analysis - e.g. up to 20 questions, using only descriptive statistics. Confirm with supervisor that the amount of work is appropriate.	Redefine scope to more appropriate levels without ensuring excessive loss of final quality of product.

Table 2: Milestones

Milestone name	Due date
Revised proposal finalised and uploaded to Vula	10/06/2019
Survey feasibility demonstration	19/07/2019
Conduct first survey and write-up	19/07/2019
Conduct final survey and write-up	29/07/2019
Write final complete draft of paper	16/08/2019
Submit final project paper	26/08/2019
Submit final survey report	02/09/2019
Final project demonstration	16/09/2019
Create project poster	23/09/2019
Create project webpage	30/09/2019
Write reflection paper	07/10/2019