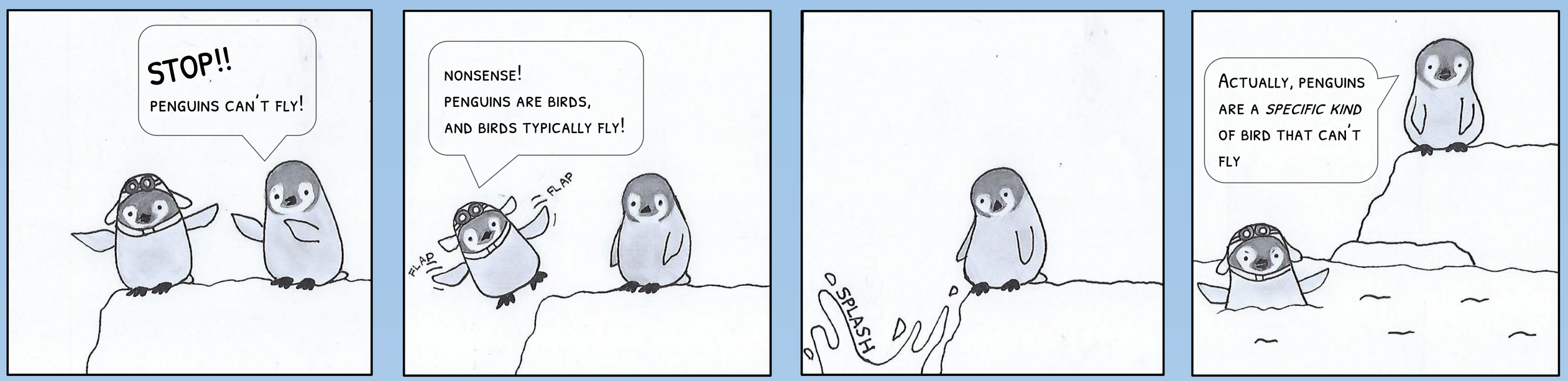
Propositional Defeasible Explanation



1 Classical Reasoning

Knowledge Representation and Reasoning expresses reasoning in terms of formal languages called *logics*. While there are many logics, we focus on *propositional logic* given its foundational role. We can encode knowledge—e.g. penguins are birds, and birds fly—in propositional logic and infer new knowledge using *entailment*:

$$\{\text{PENGUIN} \rightarrow \text{BIRD},\ \text{BIRD} \rightarrow \text{FLIES}\} \models \text{PENGUIN} \rightarrow \text{FLIES}$$

entails

2 Explanation

Sometimes we want to know why certain entailments hold. *Explanation* tells us which statements in our knowledge base are responsible for a given entailment. Suppose we know the unrelated information that birds have eyes:

$$\{\text{PENGUIN} \rightarrow \text{BIRD},\ \text{BIRD} \rightarrow \text{FLIES},\ \underline{\text{BIRD} \rightarrow \text{EYES}}\}$$

Explanation shows that only some of these statements are significant to the entailment that penguins fly. Specifically, we find that the statements

$$\{\text{PENGUIN} \rightarrow \text{BIRD},\ \text{BIRD} \rightarrow \text{FLIES}\}$$

are the only statements relevant for the conclusion.

Explanation has been shown to be an important aspect of reasoning systems. The concept is particularly useful for knowledge base debugging but has also been shown to be important for user comprehension and confidence in reasoning systems.

3 Defeasible Reasoning

Classical reasoning can be limiting because it cannot easily model statements that *typically* hold, but for which there might be exceptions. *Defeasible reasoning* allows us to reason with this uncertainty. So to model penguins a bit more accurately, we rather say:

$$\{\text{PENGUIN} \rightarrow \text{BIRD},\ \underbrace{\text{BIRD} \mathrel{|\!\sim} \text{FLIES}}_{\textit{birds typically fly}},\ \underbrace{\text{PENGUIN} \rightarrow \neg\text{FLIES}}_{\textit{penguins don't fly}}\}$$

We utilize a particular approach to defeasible reasoning called the KLM approach, focusing on three definitions for defeasible entailment: *Rational Closure*, *Relevant Closure* and *Lexicographic Closure*.

4 Defeasible Explanation

Explanation has not yet been explored for defeasible reasoning apart from some introductory work. This work includes ideas such as *weak* and *strong explanation* which are each notions of explanation for defeasible logic.

Aim: To improve our theoretical understanding of defeasible explanation and propose algorithms for evaluating defeasible explanations within the KLM framework.

Results:

- We adapt weak justification—previously only explored for Rational Closure—to Relevant and Lexicographic Closure.
- We propose an algorithm for the evaluation of strong justifications for KLM-style defeasible reasoning.
- We give a series of intuitive properties expected for sensible defeasible explanation and show that weak justification satisfies these properties.

University of Cape Town
Department of Computer Science
Email: dept@cs.uct.ac.za

IT | SCHOOL OF IT

Lloyd Everett evrllo001@myuct.ac.za
Emily Morris mrremi007@myuct.ac.za
Supervised by Prof. Thomas Meyer