Title: Research Proposal

Ryuta Arisaka[1], Francesco Santini[2]

[1] Kyoto University, Kyoto, Japan
ryutaarisaka@gmail.com
[2] Department of Mathematics and Computer Science, University of Perugia, Italy
francesco.santini@unipg.it

Abstract. ???

1 Introduction

Suppose the case of manslaughter where a man is pleaded guilty of killing his wife. Suppose there were 100 similar manslaughter cases of a man killing his wife in the last 5 years, in 60 of which the suspects were charged with voluntary manslaughter, and in 40 of which they were charged with involuntary manslaughter. Based on those past outcomes, we may model judge's decision-making process as Fig. 1's right hand side of argumentation, where $\Rightarrow$ (in any orientation) indicates a support relation and $\rightarrow$ (in any orientation) an attack relation. It intends to prescribe to the current manslaughter case the following
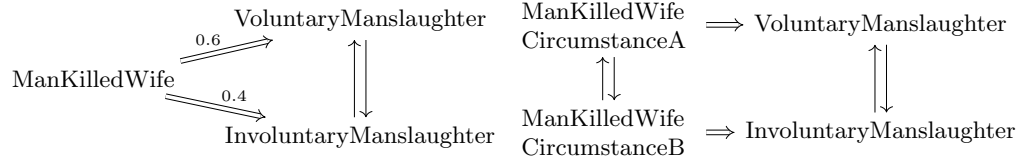

Fig. 1. Left: argumentation with disjunctive branches from acceptance of ManKilledWife. **Right**: argumentation with no disjunctive branches from acceptance of ManKilledWifeCircumstanceA or ManKilledWifeCircumstanceB.

disjunctive guideline: statistically, by 60% chance, it is a voluntary manslaughter, and by 40% chance, it is an involuntary manslaughter.

1.1 Research problems

While the statistically disjunctive argumentation provides an intuitive overview of court decisions on manslaughter cases, it is rather evident that the judge cannot base his/her judgement on the statistics for his/her decision on the present manslaughter case. Between voluntary and involuntary manslaughter, the difference in severity of penalty is in fact too huge to tolerate probabilistic judgement.

What the judge will need is more detail that can somehow split ManKilledWife argument into the argumentation on the right hand side of Fig. 1, so that if the current manslaughter case happens to fall into CircumstanceA, then the judge can deterministically conclude the case as voluntary manslaughter, and if it should fall into CircumstanceB, then the judge can deterministically conclude the case as involuntary manslaughter. Or if it does not fall into either, then the judge can, again deterministically, conclude the present case to be an exception to past decision-makings.

However, how the splitting (we call the arguments as the result of splitting 'splits') may be achieved is the primary issue coming with several associated criteria.

Mutual exclusion: Splits of an argument lead to distinct decisions.
Completeness: Splits of one argument taken together recover the original argument for the past cases.
Generality: Splits are general enough for future uses of the argumentation as a decision-making guideline.

Example 1. To illustrate each of the three criteria, let us suppose the following. The left table shows the past 100 cases and the right table shows the present case. For simplicity, there were 30 past cases all falling under the first description, 30 past cases all under the second description, and 40 past cases all under the third description.

Description	Decision
The killer stabbed the neck with a knife. Blood gushed out, and the victim was pronounced dead on the spot.	Voluntary manslaughter
The killer choked the victim who was pronounced dead hours later.	Voluntary manslaughter
The victim was shot dead on the spot. The killer is a cop chasing a burglar. A warning shot hit the victim.	Involuntary manslaugher

Description	Decision
The killer stabbed the chest with a knife multiple times. The victim was pronounced dead on the spot.	

- Suppose CircumstanceA is "the wife dead on the spot" and "the man used a knife", and CircumstanceB is "the wife dead on the spot" and "the man used a gun". Then, The first description fits in CircumstanceA, and the third description is under CircumstanceB. Since the first description's decision is voluntary manslaughter, while the third description's decision is involuntary manslaughter, we have achieved a mutually exclusive splitting. However, the second description is neither CircumstanceA nor CircumstanceB. Hence they fail to satisfy the criterion of completeness.
- Suppose CircumstanceA is "XXXX" and CircumstanceB is "YYYY".
- Suppose CircumstanceA is "ZZZZZ" and CircumstanceB is "WWWWW".

2 Related Work

2.1 Formal argumentation and ontologies

In [5] the authors present an argumentation framework for reasoning and management in (inconsistent or incoherent) description logic ontologies which contain conflicts. First, a new argumentation framework obtained by combining Besnard and Hunter's framework with binary argumentation is introduced to frame the inner relation over axioms in an ontology. A dialogue mechanism, based on this framework, is then presented to derive meaningful consequences from inconsistent ontologies. Three novel operators are developed to repair those axioms or assertions which cause inconsistency or incoherency of ontologies by using this framework.

2.2 Disjunctive frameworks

The notion of disjunctive attacks was introduced for the first time in 2009 in [2]. This works presents joint and disjunctive attacks; formally, x attacks $\{e_1, \ldots, e_m\}$ disjunctively means: $x = in$ implies $\bigwedge_{i=1}^{m} e_i = out$ (especially this can mean that if x is in then several or more e_i are out). If x is labelled in and $e_j, j = 1, \ldots, k$ are the nodes being disjunctively attacked from x then at least one of e_j must be out.

In [3] the authors discusses properties of what is called argumentation systems with disjunctive attacks, where we have an argument x attacking a set of arguments H, without specifically attacking any specific $y \in H$: in other words, x attacks the conjunction $\bigvee_{y \in H} y$. In this paper, the representation theorem turns the concept of disjunctive attacks into meta-level concept

Combined attacks has been considered in [4]. There, the authors generalise Dung's approach to allow it to handle sets of attacking arguments. However, the aim is to let the framework be capable of dealing with *synergies* among arguments in *and*, thus *without* considering numerical strengths, something that is at the core of this work instead. For instance, "a_1: *Joe does not like Jack*", and "a_2: *There is a nail in Jack's antique coffee table*" do not attack separately "a_3: *Joe did not hammer a nail into Jack's antique coffee table*", but they do it only in conjunction (example taken from [4]).

A weighted defence based on combined attacks in [4] was presented in [1], with the main goal to study well-founded frameworks by using such a new notion of defence.

3 Background

4 Conclusions and Future Work

Research problems and objectives

The current shortage of argumentation research correct me if this is not correct.... is, modelling of specific argumentations that have been already conducted is studied extensively, but reusability of the derived argumentation has been seldom shed light on. As such, while useful as a formal model to orderly understand a decision-making process, a new specific argumentation case has to be modelled anew, which is expensive and tedious. That is the same problem as model checking suffers from.

Therefore the aims:

1. We develop a formal framework which enables construction of a general argumentation from similar past decisions that can act as a guideline for a similar future decision-making case.
2.

Significance

1. We fortify XAI by having a formal mechanism to explain criteria for decision-makings.
2.

References

1. Bistarelli, S., Santini, F.: Well-foundedness in weighted argumentation frameworks. In: Logics in Artificial Intelligence - 16th European Conference, JELIA. Lecture Notes in Computer Science, vol. 11468, pp. 69–84. Springer (2019)
2. Gabbay, D.M.: Fibring argumentation frames. Studia Logica 93(2), 231–295 (2009)
3. Gabbay, D.M., Gabbay, M.: Theory of disjunctive attacks, part I. Log. J. IGPL 24(2), 186–218 (2016), https://doi.org/10.1093/jigpal/jzv032
4. Nielsen, S.H., Parsons, S.: A generalization of dung's abstract framework for argumentation: Arguing with sets of attacking arguments. In: Argumentation in Multi-Agent Systems ArgMAS. LNCS, vol. 4766, pp. 54–73. Springer (2006)
5. Zhang, X., Lin, Z.: An argumentation framework for description logic ontology reasoning and management. J. Intell. Inf. Syst. 40(3), 375–403 (2013), https://doi.org/10.1007/s10844-012-0230-7