A Puzzle about Monotonicity under Epistemic Verbs

Jialiang Yan & Maria Aloni

ILLC, University of Amsterdam

Abstract. It is controversial whether monotonicity (especially upward monotonicity) is blocked under attitude verbs expressing obligation or desire. This debate tends to focus on whether the classical semantics for these verbs is correct. The goal of this paper is to suggest a pragmatic account for the puzzles against monotonic semantics in the controversy. A similar puzzle is recreated in epistemic setting where the monotonic semantics from a center point of view is uncontroversial. We observe that there are still extra effect of monotonicity under knowledge. An analysis is proposed to explain how this effect arises. Building on [2], we develop the Quantified Bilateral State-based Epistemic Logic (QBSEL) to capture the analysis formally.

1 Introduction

Monotonic inferences are ubiquitous in natural language. The monotonicity principle is a key to comprehending the mechanisms of human reasoning. In Aristotelian logic (especially for syllogisms) or in the ancient Chinese Mohist logical thought, many valid patterns of reasoning can be understood in terms of monotonicity explicitly or implicitly, see [26,22]. In modern Natural Logic ([24,17]), monotonicity calculus is the primary technical apparatus for processing natural language reasoning. To set the scene, some basic definition should be introduced. According to [25], the monotonic step which guarantees logical validity could be named Predicate Replacement which amounts to the substitution of a specific (general) predicate with a general (specific) one in the upward (downward) context generated by the upward (downward) monotonic determiner. So the notion of monotonicity can be defined in the first-order logic as following:

Definition 1 (Mon-inf). [1]

– Upward monotonicity.

[1] In [25], there is another notion of monotonicity, that is Extension Increase, and it can be defined as Mon-sem (only focus on upward case):

$M \equiv_P^+ M'$, where M' is the same model of M except for the Interpretation of P and $I(P) \subseteq I'(P)$, then $M', v \vDash \phi(P)$

It can be proven that the Inferential monotonicity is equivalent to Semantic monotonicity.

From $\phi(P)$ and $\forall x(P(x) \to Q(x))$, it follows that $\phi(Q/P)$, where $\phi(Q/P)$ is the result of replacing each occurrence of P in ϕ by Q.

– Downward monotonicity.
From $\phi(Q)$ and $\forall x(P(x) \to Q(x))$, it follows that $\phi(P/Q)$, where $\phi(P/Q)$ is the result of replacing each occurrence of Q in ϕ by P.

The following examples will be useful to understand the above definitions.

(1) a. Some dogs bark ⇒ Some animals bark (upward)
b. Some dogs are barking loudly ⇒ Some dogs are barking (upward)

(2) a. Every student in the school is running ⇒ Every female student in the school is running (downward)
b. Every reading room is equipped with a Mac computer ⇒ Every reading room is equipped with a computer (upward)

Monotonicity describes a functional property of 'ordering preserving'. In natural language, the upward monotonicity of an operator is commonly treated as upward entailing, namely the monotonicity pattern can preserve the direction of entailment in its argument. In contrast, the downward monotonicity is then the downward entailing that reverses the entailment. The inference in (1-a) shows that the determiner 'some' preserves the entailing 'dogs ⇒ animals': 'some dogs ⇒ some animals', so it is upward monotonic in the first argument 'dogs/animals'. While the determiner 'every' in (2-a) makes the inference from 'students' to 'female students' true, which reverses the entailing 'female students ⇒ students', so it is downward monotonic in its first argument 'student/female student'.

A further fact is that negation is downward entailing, so it reverses the direction of a monotonic entailment embedded under it. For example:

(3) a. No dogs bark ⇒ No black dogs bark (downward)
b. No dogs are barking ⇒ No dogs are barking loudly (downward)

(4) a. Not every student in the school is running. ⇒ Not every person in the school is running (upward)
b. Not every reading room is equipped with a computer ⇒ Not every reading room is equipped with a Mac computer (downward)

As can be seen, monotonicity patterns in (3)-(4) are exactly the opposite to those in (1)-(2).

Likewise, if an attitude verb, which is semantically interpreted to be a propositional function, is assumed to be upward (downward) entailing, then it should preserve (reverse) the monotonic entailment in its sentential complement. However, it's not as straightforward as we expect. The monotonic reasoning under the attitude verbs expressing obligation or desire sometimes appears to be blocked, while sometimes doesn't. Therefore it is controversial to determine whether the classically monotonic semantics for deontic or bouletic verbs is correct in its prediction of monotonicity.

In this paper, we suggest a pragmatic strategy to account for the puzzles in the discussion. For this reason, we first go and look in other domains where monotonic semantic account is uncontroversial. We will discuss cases of apparent monotonicity failure under epistemic verbs. In doing so, we expect to provide a credible analysis to explain why monotonic inferences under attitude verbs sometimes appear to fail.

The article is structured as follows: In Sec.2 we first present a puzzle to show that monotonicity under knowledge will introduce problems in practice. While the examples shown later will demonstrate that such trouble sometimes disappears. We will discuss the causes of the difference. In Sec.3 and Sec.4, we propose the core analysis to account for these cases, which is based on an aspect of the pragmatics of linguistic communication. To explain how some key reasoning is derived, we develop the framework Quantified Bilateral State-based Epistemic Logic based on [2]. And hence we provide a formal account to capture the pragmatic effects. Finally, we conclude in Sec.5.

2 The controversial issue about monotonicity

As introduced in Sec.1, the issue about monotonicity becomes important to determine whether the classically monotonic semantics for deontic and bouletic verbs is correct. Both sides of the debate present plausible examples and reasonable arguments to illustrate their position. In this section, we first present some typical examples to give an overview of the arguments. Then we explain our motivation to study the phenomenon in epistemic domain.

2.1 Non-monotonic v.s. monotonic semantics

There exist prima facie examples showing that monotonic reasoning is blocked under deontic and bouletic modalities. Some people argued that the classical semantics for them is incorrect in its prediction of monotonicity, and hence some non-monotonic alternatives are required (such like [11,16,6,15].). The following counterexamples are famous for providing challenges to the classical monotonic semantics.

(5) a. Nicholas wants to get a free trip on the Concorde $\not\Rightarrow$
    b. Nicholas wants to get a trip on the Concorde. ([4,11])

(6) a. It is obligatory that the letter is mailed.
    b. It is obligatory that the letter is mailed or burned. ([20])

An oddness can be observed in (5): the desire of wanting a free trip for an agent cannot guarantee her desire of any trip. Most of us would probably not want a trip pn the Concorde unless it is a free one. The inference in (6) is well-known as Ross' Paradox, where the obligation to mail the letter or burn it is derived from the obligation to mail the letter. In other words, the obligation of mailing the letter can be fulfilled by some forbidden action (burning the letter), which is rather strange. As a result, there doesn't seem to be an upward (or downward) entailment in (5) or (6). There are plenty such examples on the market against monotonic semantics for the deontic and bouletic modalities.

On the other hand, there are some cases where the monotonicity seems to be required. For instance:

(7) a. You don't have to bring any alcohol to the party.

(8) a. Jon doesn't want to meet any student.

(reproduced from [27])

In both cases, if we interpret 'have to' and 'want' non-monotonically, then the proposition 'You have to bring alcohol to the party' and 'Jon wants to meet student' will not be the upward entailing. Then the negation 'don't have to' and 'don't want to' is not downward monotonic. So there is no downward environment in(7) and (8), which is necessary for licensing the NPI 'any' by the standard theory for NPIs.

It seems hard to have an argument that is completely convincing. For there are always examples popping up which give you the opposite intuition. For this reason, we are motivated to look at other domains where the classically monotonic semantics is less controversial.

2.2 Study monotonicity under knowledge

In order to clarify the causes of the monotonicity failure under attitude verbs, we are motivated to look in the epistemic domain where the classically monotonic semantics for knowledge is less problematic, so it's normally accepted.

According to canonical interpretation of epistemic modalities ([12]), "know" behaves like the necessity modality $\Box$. So at a possible world w the proposition "an agent a knows ϕ" is true if and only if ϕ holds in every epistemic possibility of a, viz., in every possible world that is epistemically accessible (indistinguishable) from w. As a result, if $\phi \rightarrow \psi$ and a knows ϕ, we can conclude a knows ψ, which is also known as the problem of Logical Omniscience. There is a lot of discussion on this problem in philosophy and logic, see [8,13]. We don't want to discuss full details here; we simply use this result to show that the semantic validity of monotonic reasoning under knowledge is guaranteed according to the canonical view, unlike in the context of desire. By this interpretation, the monotonic inferences under the verb "know" seems reliable, such like:

(9) a. Tom knows Susan is a wealthy lady
    b. Tom knows Susan is a lady

It is easy to check the validity of (9) in the canonical analysis "Susan is a lady" holds in every possible worlds where "Susan is a wealthy lady" is true. It is uncontroversial to say the inference (9) is sound.

But now let's set some conditions in the context: (i) Tom wants to marry a wealthy woman; (ii) suppose you only know the conclusion (9-b), and of course you're familiar with the context (i). Will you believe that Tom will marry Susan in this situation? People probably answer no, because they are not sure whether Tom knows Susan is wealthy. And perhaps Tom thought Susan was a poor lady. Our analysis of this case is that the inference from (9-a) to (9-b) leads to a loss of information and because of this some additional inferences can be derived pragmatically from (9-b). In this example, the context (i) and (ii) combining with the conclusion (9-b) will make people think that Tom might believe that Susan is poor, and hence it will affect their prediction to the question.

In the next section, we will present refined examples to show that even under the epistemic verbs, whose classically monotonic semantics is less problematic, there will be potential failure of monotonicity.

3 Monotonic inferences under epistemic verbs

3.1 A puzzle

Consider the following example:

> Example 1
>
> Tom is the organiser of a pet party at ILLC, which is open to all kinds of pets. He thinks Susan may need to be informed that cats are at this party as she is allergic to cats. Meanwhile the drinks for the party have not yet arrived so he needs to pick them up. For some reason Tom can't do both at the same time, and he has to decide which one to do. Tom is torn at the moment because either Susan has a bad allergic reaction from the party or the lack of drinks will make the party fail.
>
> Just then Tom hears two colleagues A and B in the common room chatting about Susan, and A says:
>
> (10) Susan knows that some animals will be at the party.
>
> Based on this piece of information Tom concludes that it is not necessary to warn Susan. Eventually, however, Susan comes to the party and gets a serious allergic reaction. Tom blames A and asks furiously "Why did you say that Susan knew that there would be animals at the party?" And A is aggrieved and replies:
>
> > What I said was true, Susan knew that animals would be at the party, because she knew there would be dogs (she saw some other colleague preparing her dog and thought that it was a dog party). So by monotonicity, I reasoned:
>
> What I said was true, Susan knew that animals would be at the party, because I knew that she knew there would be dogs. So by monotonicity, I conclude (10) from:
>
> (11) a. Susan knows that some dogs will be at the party.
> b. Susan knows that dogs are animals.
>
> It seems that A said something true after all. But relying on a true statement Tom made the wrong decision, one with disastrous consequences for his friend. What went wrong?

Was A's reasoning incorrect after all? If not, how did Tom arrive at such a bad decision based on a true statement? Was Tom's decision to go to supermarket completely wrong? In what follows I will argue that (i) we need monotonicity reasoning under knowledge and (ii) such reasoning can be useful for practical decision making, but also harmful, as shown in Example 1, depending on contextual factors. The goal of the following section is to understand what these contextual factors amount to.

3.2 The impact of context

Consider the following scenario:

> Example 2:
> Tom has to take a zoom exam and is worried that he will fail it, so he asks his friend Susan to help him.
>
> There are now two final questions left, and if both are answered correctly, Tom will pass the exam. Suppose Susan has only one opportunity to pass the answer. The first key question asks if there exist mammals that depend on the ocean for existence. Susan has to decide whether to pass the answer to Tom, and she knows the following two facts hold true:
>
> (12) a. Tom knows that some whales rely on the ocean for existence.
> b. Tom knows that whales are mammals.
>
> By (12) and monotonicity, Susan concludes that
>
> (13) Tom knows that some mammals rely on the ocean for existence.
>
> And hence, Susan predicts that Tom knows the answer to the question. So she doesn't act. Susan makes a right decision because Tom answers correctly to the mammal question and Susan still has a chance to pass the answer to the last question, which happens to be something Tom doesn't know. So Tom eventually passes the exam!

This example shows that monotonicity can be useful and also lead to correct decisions in epistemic scenarios, which was not the case in the scenario described in Example 1. What is the difference between Example 1 and 2? We will argue that the crucial difference lies in the issue under discussion. Consider the following variant of Example 1:

> Example 3:
> Assume the same scenario as in Example 1 and further suppose that the two colleagues A and B in the common room were having a bet on whether Susan would answer correctly to the following question:
>
> (14) Would there be animals at the ILLC party?
>
> In the described scenario it would be perfectly rational for A to bet that Susan would answer yes according to the reasoning in (11).

In this example, the monotonic inference (11) helps A make to a correct prediction and a right decision. Why doesn't monotonic reasoning lead to the right decision in Example 1, while it does in this example?

We argue that the context plays a very important role, and the crucial contextual aspect is the issue under discussion, i.e., the question that needs to be answered to arrive at a solution to the current decision problem. It is clear that in Example 3 the issue under discussion is "whether Susan knows that some animals will be at the party", a question about the predicate "animals". The information from (11) is sufficient for the colleague A to make the right decision, because the conclusion appropriately settles the issues under discussion. Likewise, the question under discussion in Example 2 is about "mammals" and the monotonic reasoning there leading to a conclusion about mammals helps Susan predict correctly the relevant facts about Tom's knowledge, and then make a right decision. In contrast in Example 1 the central issue is "whether Susan knows that some cats will be at the party" which makes the predicate "cats" salient, rather than "animals". So statement (10), which is derived from (11) communicating information about animals, cannot support Tom in making correct predictions on knowledge about cats. It is only in this latter case that the loss of information caused by the monotonicity step becomes problematic.

One question that arises though is whether Tom was completely wrong in concluding from (10) that Susan didn't need to be warned. Tom was clearly wrong, but it was not completely irrational for him to assume that Susan would not come to the party from the fact that she knew that animals would be there. After all cats are animals, and Susan knew herself to be allergic to cats so she would be more sensitive to cat-related matters. Therefore it is reasonable to assume that she would not join the party if she aware of the possibility of cats being there. It seems that from (10) Tom concluded that

(15) Susan knows that there might be some cats at the party.

Based on (15), Tom decided not to warn Susan. Why did Tom arrive at this conclusion? In other words, was the process by which he arrived at (15) rational? We think so. As we will argue in the next section, Tom reached this conclusion by a pragmatic inference which is similar to the ignorance inferences triggered by disjunctive statements. Consider the following example:

(16) a. Jack bought a Porsche or a Ferrari $\rightsquigarrow$
b. The speaker does not know which car Jack actually bought. It might be a Porsche and it might be a Ferrari.

This example shows the use of a disjunction can convey the inference that the speaker does not know which disjunct is true in a conversation. Not only the ignorance inference, but the phenomenon of free choice inferences (FCIs) also shows that a conjunction can be derived from a disjunctive statement (see [14,3,29]):

(17) a. You may ask the office for a laptop or a desktop $\rightsquigarrow$

b. You may ask the office for a laptop and you may ask the office for a desktop.

Both the premise (16-a) and (17-a) convey a disjunctive meaning from which a conjunctive meaning expressed in the conclusions (16-b) and (17-b) derived, which is against the prescriptions of classical logic yet is rational to reach it. A possible account for these phenomena resorts to theories of pragmatics. In this way, as argued in [19], the premises of inferences from a disjunction can be interpreted in a non-obvious but predictable and reasonable way. So the plausibility and rationality of these inferences can be reflected in the pragmatic aspect as will be discussed in the next section.

4 Ignorance and disjunctive analysis of predicates

4.1 Ignorance inferences out of disjunctions

Before presenting our analysis, it is necessary to discuss more in details the ignorance effects that disjunctions can trigger when used in ordinary conversations, as in (16) repeated as (18) in the following:

(18) a. Jack bought a Porsche or a Ferrari $\rightsquigarrow$
b. The speaker does not know which car Jack actually bought. It might be a Porsche and it might be a Ferrari.

As a possible mechanism accounting for this phenomenon, this effect can be derived from Grice's cooperative principle as an application of his Maxim of Quantity ([10]):

Maxim of Quantity

- Make your contribution as informative as is required (for the current purposes of the exchange).
- Do not make your contribution more informative than is required.

According to Gricean reasoning, upon hearing statements like (16-a), the recipients will draw an inference that (i) the speaker knows that Jack bought a Porsche or Ferrari, but (ii) she doesn't know exactly which car Jack bought. Ignorance effects are generated from disjunctive statements with above principle. Because if the speaker had known one of the disjuncts, she would have asserted it as true to make the conversation informative according to Maxim of Quantity, rather than using a disjunction to describe the situation. On Grice's account, when a disjunctive statement is used, it conveys that the speaker has not enough evidence to support one disjunct and negate the other.

There are a lot of discussions on ignorance inferences, e.g. [9,18,5,21]. For the epistemic meaning of a disjunction $A \vee B$, as mentioned, the speaker is not certain if A and B are true, so the epistemic possibility of each disjunct is open to the recipients. In other words, it can be interpreted as conjunctions of epistemic possibilities, for example, $A \vee B$ means A might be the case and B might be the case, from which we have:

(19) $A \vee B \rightsquigarrow \Diamond A \wedge \Diamond B$

where $\Diamond$ represents the epistemic possibility "might" according to the speaker.

In this paper, however, we derive ignorance inferences (and also other modal inferences triggered by disjunctive statements like FCIs) in a different way. Following [1], different from Gricean or neo-Gricean approaches, Aloni has provided a logic-based account in which assuming that people prefer creating a vivid picture representing the world while understanding a statement, and hence they systematically neglect abstract situation that validate sentences by virtue of some empty configuration. Ignorance effects and related inferences can be derived from this assumption ([1,2]). How this framework works will be explained in Sec. 4. In the coming subsection, the phenomenon of ignorance inferences will be employed to analyze Tom's reasoning behavior in Example 1.

4.2 Disjunctive meaning of the predicate

In this section, we will address the question why Tom drew the following inferences:

(20) a. Susan knows some animals will be at the party $\rightsquigarrow$ (10)
b. Susan knows that some cats might be at the party (15)

We argue that the status of the inference from (20-a) to (20-b) is a pragmatic inference which is similar to the ignorance inference triggered by disjunctive statements. At the core of our proposal is the assumption that in certain contexts a predicate can express a disjunctive meaning.

In a given situation, the issue under discussion can make a predicate salient. For example, when a kid who wants chocolate is shopping with her mother, in this situation she is really care about the question that "will my mom buy chocolate for me?". The question under discussion makes the predicate "chocolate" salient. Then if her mother says "We're going to buy some snacks", the kid will wonder with great excitement "Are we going to buy chocolate?". In this case, the predicate "snacks" is naturally reinterpret as the disjunction "chocolate or any other snacks that are not chocolate".

Semantically, in the domain the predicate "snacks" is interpreted as a disjoint union of sets of "chocolate" and "snacks that are not chocolate". In other words, the predicate "snacks" has been partitioned by classifying snacks into chocolate and those that are not (see Fig 1 below). So we propose that when a predicate Q which semantically includes P ($P \subset Q$) is uttered and P is salient, Q will be understood with respect to a partition provided by P and its negation. We will call P the sub-predicate of Q. In the example,"chocolate" is exactly the sub-predicate of "snacks".

As the case in Example 1, Tom focuses on the question if Susan knows that some cats will be at this party, and hence the predicate "cats" becomes salient in this context. In this situation the meaning of the predicate "animals" is partitioned into the sub-predicates "cats" and "animals that are not cats", and hence

Fig. 1. Partition of snacks

"animals" is reinterpreted by Tom as "cats or animals that are not cats". From now on we implicitly assume that "non-cats" means animals which are not cats. Therefore the statement (10) now can be interpreted as:

(21) a. Susan knows some animals will be at the party $\Longleftrightarrow$
b. Susan knows that some cats or non-cats will be at the party.

According to (19) mentioned in Sec 3.1, the disjunctive meaning of the predicate "animals" will give rise to an ignorance effect:

(22) a. Susan knows that some cats or non-cats will be at the party $\rightsquigarrow$
b. Susan knows that there might be some cats at the party, and knows that there might be not cats at the party.

Therefore, from (22), Tom concludes that "Susan knows that there might be some cats at the party" stated by (20-b). This is the process how Tom gets his prediction, which is incorrect but rational since it follows from logic and pragmatics.

We now seem to be in a position to address the question: why Tom's behaviour is rational. To illustrate this point, we propose that in a given context a predicate can convey a disjunctive meaning, from which an ignorance effect is generated and licenses to predict that "Susan knows that there might be some cats at the party". Relying on this possibility, Tom makes the wrong decision.

So far we conclude that some modal inferences will be triggered by the disjunctive statements conveyed from the monotonic reasoning. In the next section, we will explain how these inferences are derived by employing a logic-based account. We will formalise this reasoning process by the framework BSML presented in [1] in which the pragmatic enrichment function (denoted by $^+$) is the main ingredient to process the pragmatic parts in the inferences.

5 Quantified bilateral state-based epistemic logic (QBSEL)

In this section, we introduce the bilateral state-based modal logic and its epistemic extension which we will use to provide a formal account of the puzzle we presented in Sec.2.1. To capture the analysis of salient sub-predicate on our proposal, we will define the disjunctive representation of a predicate. If $Q = P \vee \neg P$, then $\lambda x(Px \vee \neg Px)$ is the disjunctive representation of Q. It is worth nothing that in classical logic, $\forall x Qx$ and $\forall x((Px \vee \neg Px) \wedge Qx)$ are logically equivalent. However in the logic introduced in this section, classically logically equivalent formulas can give rise to different pragmatic effects due to possible pragmatic intrusion. We will exploit this formal analysis to account for the differences between the examples discussed before. Example 1 will be different than 2 & 3 because in 1 the predicate is reinterpreted as its disjunctive representation.

5.1 The basic elements in BSML

As mentioned in Sec 3.1, ignorance inferences out of disjunctive statements can be analyzed in term of pragmatics. Gricean or neo-Gricean approaches have provided a possible explanation which appeals to some conversational reasoning, in which semantics and pragmatics are modeled as two separate components.

In [1], Aloni proposed a framework, the bilateral state-based modal logic (BSML). The idea underlying the framework is that modal inferences such like the free choice and ignorance inferences are not the outcome of reasoning, but are direct consequences of something else that the speaker do in a conversation. People create conversational representations of the world while understanding a statement, and in doing so, they systematically neglect models that validate sentences by virtue of some empty configuration (zero-model). In other words, people prefer to imagine a concrete reality rather than an abstract one. Aloni called this tendency neglect-zero. The following example shows what the zero-models are for the statement "Every bowl has an apple":

Statement: Every bowl has an apple

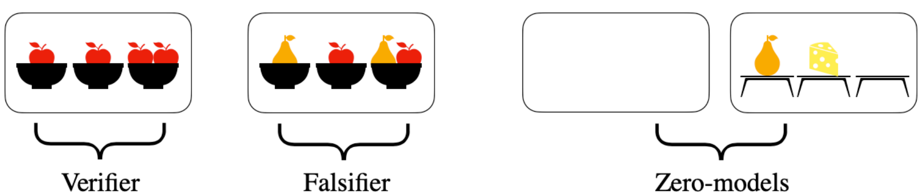

Fig. 2. Example of zero-models

The first picture shows a model which verifies this statement, and the second doesn't. Zero-models are the blank picture or the picture that describes something completely unrelated to apples. To describe the representation of the world (or a picture of the partial reality) created by people, a bilateral state-based framework BSML is adopted, in which the world is represented by information states. And the assertability and rejectability of a sentence are modeled by support ($\vDash$) and anti-support ($\Dashv$) conditions respectively in the model, rather than the truth. In [1], Aloni showed that the free choice inferences directly follow from neglect zero effects in the propositional version of BSML. Aloni and van Ormondt ([2]) developed first-order version of BSML, a quantified bilateral state-based modal logic (QBSML), to interpret modified numerals and some related puzzles.

The core idea of BSML is providing a method to formalize the neglect-zero effects by employing the non-emptiness atom (NE) from team logic ([28]). The atom NE is supported by a state iff the state is non-empty.

$$\mathcal{M}, s \vDash \mathrm{NE} \;\text{ iff }\; s \neq \varnothing$$

In fact, the atom NE also connects to the principle "avoid contradictions" in conversations. In state-based semantics, the empty state can vacuously supports every classical formula, including contradictions. The interpretation of NE in BSML requires the supporting state to be non-empty. And hence, it captures the idea that the empty states which represent abstract elements (zero or empty set in mathematics) and support logical insanity is excluded in the process of human cognition.

Then a pragmatic enrichment function (denoted by $[\cdot]^+$) can be defined in terms of a systematic intrusion of NE in the process of interpretation in [1]. The following definition is based on the language of QBSML proposed by [2]:

Definition 2 (Pragmatic enrichment). A pragmatic enrichment function is a mapping $[\cdot]^+$ from the NE-free fragment of language $\mathcal{L}$ in QBSML to $\mathcal{L}$ such that:

- $[Pt_1...t_n]^+ = Pt_1...t_n \wedge \mathrm{NE}$
- $[\neg\phi]^+ = \neg[\phi]^+ \wedge \mathrm{NE}$
- $[\phi \vee \psi]^+ = ([\phi]^+ \vee [\psi]^+) \wedge \mathrm{NE}$
- $[\phi \wedge \psi]^+ = ([\phi]^+ \wedge [\psi]^+) \wedge \mathrm{NE}$
- $[\Diamond\phi]^+ = \Diamond[\phi]^+ \wedge \mathrm{NE}$
- $[\exists x\phi]^+ = \exists x[\phi]^+ \wedge \mathrm{NE}$

Applying the pragmatic enrichment function to a disjunction will generate non-trivial effects such as the prediction of FCIs and ignorance.

It is worth nothing that the notion of disjunction in BSML is different from the connective in classical logic, which is from team logic and dependence logic ([23,28]), called split or tensor disjunction. A split disjunction $\phi \vee \psi$ is supported with respect to a state s iff s is the union of two substates each supporting one of the disjuncts:

$$\mathcal{M}, s \vDash \phi \vee \psi \text{ iff } \textit{there are } t, t' : t \cup t' = s \text{ and } \mathcal{M}, t \vDash \phi \text{ and } \mathcal{M}, t' \vDash \psi.$$

In this way, $[\phi \vee \psi]^+$ is supported by a state s iff s can be split into two non-empty substates, each supporting one of the disjuncts.

With these basic notions in the framework BSML, the modal inferences mentioned above can be captured. In order to formally discuss the puzzle we present in previous sections, in the next section we extend the first-order version of BSML to a epistemic framework quantified bilateral state-based epistemic logic (QBSEL).

5.2 Epistemic framework for BSML

The language $\mathcal{L}_\mathcal{E}$ for QBSEL is the first-order language enriched with the non-emptiness atom NE and epistemic operators.

Definition 3 (Language $\mathcal{L}_\mathcal{E}$). The language $\mathcal{L}_\mathcal{E}$ of Quantified Bilateral State-based Epistemic Logic is defined as following in BNF:

Term $t ::= c \mid x$

Formula $\varphi ::= P^n t_1 ... t_n \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \exists x\varphi \mid K_a\varphi \mid \Diamond \varphi \mid NE$

The predicate constants $P^n \in \mathcal{P}^n$, $n \in \mathbb{N}$, the individual constants $c \in \mathcal{C}$ and the variables $x \in \mathcal{V}$. The formula $K_a\varphi$ is read as "the agent a knows that φ". The operator $\Diamond$ represents the epistemic possibility, and $\Diamond\varphi$ is read as "it might be that φ".

Definition 4 (Epistemic Model $\mathcal{M}$).

An epistemic model for $\mathcal{L}_\mathcal{E}$ is a tuple $\mathcal{M} =< W, D, R, I >$:

- W is a set of possible worlds.
- D is a non-empty set which is assumed to be constant across worlds.
- R is an accessibility relation on W, which is reflective, transitive and symmetric. We define the following abbreviation:

  $R(w_i) = \{v \in W \mid w_i R v\}$. .
- I: $W \times \mathcal{C} \cup \mathcal{P}^n \to D \cup \mathscr{P}(D^n)$ is an interpretation function which assigns entities to individual constants and sets of entities to predicate letters relative to worlds $w \in W$:

  $I(w)(\gamma) = \begin{cases} d \in D & if\ \gamma \in \mathcal{C} \\ S^n \subset D^n & if\ \gamma \in \mathcal{P}^n \end{cases}$

As mentioned, BSML is built on state-based semantics in which the formulas are interpreted with respect to information states. An information state in the first-order modal framework is defined as a set of indices. An index i is a pair i=$< w_i, g_i >$ where $w_i \in W$ and $g_i = \mathcal{V} \to D$. By the indices, an information state can be encoded information about the value of variables in worlds. For all indices i, k in a state s, the assignments g_i and g_k have the same domain. Having an information state, we can define what is called knowledge state:

Definition 5 (Knowledge state).

$s_i^K := \{< w_{ij}, g_i > | < w_i, g_i > \in s, \text{ and } w_{ij} \in R(w_i)\}$

We define a knowledge state to represent agent's epistemic states. A toy example has been shown in Figure 3, where the knowledge states are represented by red circle and we assume the assignment constant. The underlying idea of this concept is that the relation R used for interpretation of knowledge and $R(w)$ models the knowledge state of the relevant agent. Notice that this might diverge from the information state s of the speaker, with respect to which we will interpret our formulas.

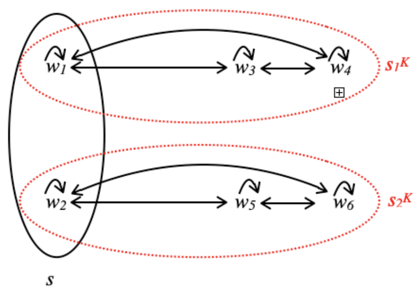

Fig. 3. Knowledge state

Three ways of extensions on states can be defined. Before that we need to introduce some notions:

Definition 6.

a. $g[x/d] := (g \backslash \{x, g(x)\}) \cup \{x, d\}$
b. $i[x/d] := w_i, g_i[x/d]$.

If the variable $x \notin \text{dom}(g)$, then it can be added and set to the value d as $g[x/d]$. An index $i[x/d]$ is the result of adding x assigned to d to $\text{dom}(g)$. Notice that if x is already in $\text{dom}(g)$, then the operation $g[x/d]$ just indicate resetting the value of x.

Based on these notions, we can define three ways of x-extension of a state s (see[7]):

Definition 7 (x-Extensions of a state s).

1. Individual x-extension of s.
   $s[x/d]$:=$\{i[x/d] \mid i \in s\}$, for some $d \in D$.
2. Universal x-extension of s.
   $s[x]$:=$\{i[x/d] \mid i \in s, \text{ and } d \in D\}$.

3. Functional x-extension of s.
   $s[x/h]{:=}\{i[x/d] \mid i \in s, \text{ and } d \in h(i)\}$, for some function $h: s \to \mathscr{P}(D)\backslash\varnothing$.

The individual and universal x-extension of s, are the states which result by extending s with the assignment $g[x/d]$ for some $d \in D$ and $g[x/d]$ for all $d \in D$ respectively. And from the definition, it follows that the individual and universal extensions are examples of functional extensions in which any state t where for each index $i \in s$ there exists an index $j \in t$ such that $j = i[x/d]$ for some $d \in D$.

An interpretation of a term in a possibility can be defined as following:

Definition 8 (Interpretation of a term t).

$$[\![t]\!]_{\mathcal{M},i} = \begin{cases} g_i(t) & if\ t \in \mathcal{V} \\ I(w_i)(t) & if\ t \in \mathcal{C} \end{cases}$$

Definition 9 (Semantics). Let $\mathcal{M}$ be a model for language $\mathcal{L}_{\mathcal{E}}$ and s be a state, we define what it means for a formula ϕ to be supported or anti-supported at s.

$\mathcal{M}, s \vDash P^n t_1 ... t_n$ iff $\forall i \in s : [\![t_1]\!]_{\mathcal{M},i}, ..., [\![t_n]\!]_{\mathcal{M},i} \in I(w_i)(P^n)$

$\mathcal{M}, s \dashv P^n t_1 ... t_n$ iff $\forall i \in s : [\![t_1]\!]_{\mathcal{M},i}, ..., [\![t_n]\!]_{\mathcal{M},i} \notin I(w_i)(P^n)$

$\mathcal{M}, s \vDash \neg\phi$ iff $\mathcal{M}, s \dashv \phi$

$\mathcal{M}, s \dashv \neg\phi$ iff $\mathcal{M}, s \vDash \phi$

$\mathcal{M}, s \vDash \phi \vee \psi$ iff $\exists t, t' : t \cup t' = s$ and $\mathcal{M}, t \vDash \phi$ and $\mathcal{M}, t' \vDash \psi$

$\mathcal{M}, s \dashv \phi \vee \psi$ iff $\mathcal{M}, s \dashv \phi$ and $\mathcal{M}, s \dashv \psi$

$\mathcal{M}, s \vDash \phi \wedge \psi$ iff $\mathcal{M}, s \vDash \phi$ and $\mathcal{M}, s \vDash \psi$

$\mathcal{M}, s \dashv \phi \wedge \psi$ iff $\exists t, t' : t \cup t' = s$ and $\mathcal{M}, t \dashv \phi$ and $\mathcal{M}, t' \dashv \psi$

$\mathcal{M}, s \vDash \mathrm{NE}$ iff $s \neq \varnothing$

$\mathcal{M}, s \dashv \mathrm{NE}$ iff $s = \varnothing$

$\mathcal{M}, s \vDash \exists x\phi$ iff $\mathcal{M}, s[x/h] \vDash \phi$, for some function $h: s \to \mathscr{P}(D)\backslash\varnothing$

$\mathcal{M}, s \dashv \exists x\phi$ iff $\mathcal{M}, s[x] \dashv \phi$

$\mathcal{M}, s \vDash K\phi$ iff $\forall i \in s$, $\mathcal{M}, s_i^K \vDash \phi$

$\mathcal{M}, s \dashv K\phi$ iff $\forall i \in s$, there is a non-empty $t_i^K \subseteq s_i^K$ such that $\mathcal{M}, t_i^K \dashv \phi$

$\mathcal{M}, s \vDash \Diamond\phi$ iff $\exists s' \subseteq s$ and $s' \neq \varnothing$ s.t. $\mathcal{M}, s' \vDash \phi$

$\mathcal{M}, s \dashv \Diamond\phi$ iff $\mathcal{M}, s \dashv \phi$.

The final two items in the above definition show the semantics of the epistemic possibility "might" says an information state s supports $\Diamond\phi$ if and only if the current information the speaker holds provides a possibility of ϕ. In the next section, we will show some results on this definition.

5.3 Some results on "might"

Let us check the semantics of duality and double-$\Diamond$ first.

- $\mathcal{M}, s \vDash \neg\Diamond\neg\phi$
  iff $\mathcal{M}, s \dashv \Diamond\neg\phi$
  iff $\mathcal{M}, s \dashv \neg\phi$
  iff $\mathcal{M}, s \vDash \phi$
- $\mathcal{M}, s \dashv \neg\Diamond\neg\phi$
  iff $\mathcal{M}, s \vDash \Diamond\neg\phi$
  iff $\exists s' \subseteq s$ and $s' \neq \varnothing$ s.t. $\mathcal{M}, s' \vDash \neg\phi$
  iff $\mathcal{M}, s' \dashv \phi$ for some non-empty subset s' of s
- $\mathcal{M}, s \dashv \Diamond\Diamond\phi$
  iff $\forall s' \subseteq s$ and $s' \neq \varnothing$, $\mathcal{M}, s' \dashv \Diamond\phi$
  iff $\forall s'' \subseteq s'$ and $s'' \neq \varnothing$, $\mathcal{M}, s'' \dashv \phi$

Then to consider the free choice under negation, we need to check the following inference(where ϕ and ψ are NE-free formulas):

- $[\neg\Diamond(\phi \vee \psi)]^+ \vDash \neg\Diamond\phi \wedge \neg\Diamond\psi$

  Proof. Suppose $\mathcal{M}, s \vDash \neg\Diamond[(\phi \vee \psi)]^+$. It follows that s is non-empty as $s \vDash NE$ and $\mathcal{M}, s \dashv [\Diamond(\phi \vee \psi)]^+$, and hence $\mathcal{M}, s \dashv [(\phi \vee \psi)]^+$ and $s \neq \varnothing$. Then we conclude that $\mathcal{M}, s \dashv \phi \wedge NE$, and $\mathcal{M}, s \dashv \psi \wedge NE$. For the former case, it follows that $\exists t, t' : t \cup t'$ and $t \dashv \phi$ and $t' \dashv NE$. Since $s \neq \varnothing$, t' is empty and $t = s$. Therefore, we have $\mathcal{M}, s \dashv \phi$, and hence $\mathcal{M}, s \dashv \Diamond\phi$. As a conclusion, we get $\mathcal{M}, s \vDash \neg\Diamond\phi$.
  It is similar to prove the case for ψ. □

Compared with definition of epistemic possibility in [2], we give a non-relation semantics of "might" but which can predict exactly the behaviour similar to Epistemic Contradiction while preserving the non-factivity of the $\Diamond$ ($\Diamond\phi \vDash \phi$).

Fact 1 (Quasi-epistemic contradiction).

$\exists x(Px \wedge \Diamond\neg Px) \vDash_{\mathcal{M}_E} \bot$

Proof. Suppose $\mathcal{M}, s \vDash \exists x(Px \wedge \Diamond\neg Px)$. It follows that (i) $\mathcal{M}, s[x/h] \vDash Px$ and (ii) $\mathcal{M}, s[x/h] \vDash \Diamond\neg Px$ for some function $h: \ s \to \mathscr{P}(D) \backslash \varnothing$. Following (i), we derive $\forall i \in s[x/h], g_i(x/d) \in I(w_i)(P)$ and for all $d \in h(i)$. But from (ii), we derive that there is a non-empty subset $t \subseteq s[x/h]$ such that $\mathcal{M}, t \dashv Px$, namely $\forall i \in t, g_i(x/d) \notin I(w_i)(P)$ for $d \in h(i)$. Since $t \subseteq s[x/h]$, (ii) shows that there there exists $i \in s[x/h]$ s.t. $g_i(x/d) \notin I(w_i)(P)$, which contradicts with (i). Therefore $\exists x(Px \wedge \Diamond\neg Px) \vDash_{\mathcal{M}_E} \bot$. □

The non-factivity of the $\Diamond$ can be proved straightforward by the definition.

However, unfortunately, by the non-relation semantics of epistemic possibility $\Diamond$ we cannot show the validity of $K\phi \vDash \phi$ which is a principle of knowledge that we want to get and it is generally valid in epistemic logic. Because the induction on ϕ is blocked when ϕ has the form of $\Diamond\psi$. In other words, the following inference is not valid in QBSEL.

Fact 2.

$K\Diamond\psi \vDash \Diamond\psi$

One can find a counterexample like Fig 4, where ψ is supported by a non-empty substate consist of w_3 and w_4, but is rejected by the state s. A possible solution to this problem is that we substitute the K operator with the doxastic one. A possibility provided by an information state may be not sufficient to form the knowledge for an agent. If the possibility is known, then, according to the canonical interpretation of knowledge, it should hold in every epistemic state (or in our terminology, in every knowledge state). However, in our framework, the epistemic possibility only represents the possibility based on one piece of information, namely on the basis of the current information state s. And the possible worlds in s are just part of all the agent's epistemic states. If the possibility is supported by s, then it only hold in some epistemic states, not all of them. So it is not justified. It probably form a belief rather than knowledge for the agent. This bring us to consider further what is the exact meaning of knowledge based on information state. We left this as an open problem to readers.

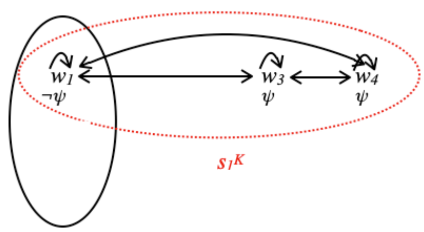

Fig. 4. Counter-example of factivity of knowledg

5.4 Revisit the puzzle

Recall our proposal that in a given context, a predicate can conveys a disjunctive meaning. We give a more refined definition of this idea in our framework. In classical logic, Qx and $(Px \vee \neg Px) \wedge Qx$ are logically equivalent. However in our logic, classically logically equivalent formulas can give rise to different pragmatic effects under the pragmatic intrusion function $[\cdot]+$, i.e. it is possible to have a couterexample where $\exists x Qx$ is supported but $[(Px \vee \neg Px) \wedge Qx]^+$ is not. So we highlight the following hypothesis in our proposal in which a predicate and the disjunction of its sub-predicates are interchangeable given a context:

Hypothesis 1 (Disjunctive representation of predicate). If P is salient in the context c, and we have

i. $P \subset Q$
ii. $Q = \lambda x(Qx \wedge Px) \vee (Qx \wedge \neg Px)$

Notice that we say P is the sub-predicate of Q only if the denotation of P is the proper subset of the denotation of Q. In addition, from now, by item (ii), we write Q as $\lambda x(P_Q x \vee \neg P_Q x)$. This condition is satisfied only if $\lambda x \neg Px$ is restricted to denotation of Q, so we express this restriction with a subscript Q. Then for every NE-free formula φ, $\varphi(Qx)$ is interchangeable with $\varphi(P_Q x \vee \neg P_Q x)$ by replacing Q in φ by $(\lambda x P_Q x \vee \neg P_Q x)$. By Hypothesis 1, the statement (19) can be formalized as:

(23) $$K_S \exists x(Ax \wedge Px) \vDash K_S \exists x((C_A x \vee \neg C_A x) \wedge Px)$$

where A and C represent the predicate "animals" and "cats" respectively. And the statement (20) can be formalized as:

(24) $$K_S \exists x[(C_A x \vee \neg C_A x) \wedge Px]^+ \vDash K_S \exists x \Diamond (C_A x \wedge Px) \wedge K_S \exists x \Diamond (\neg C_A x \wedge Px)$$

To check the following inference, we can explain how does the ignorance inference in (24) happen by our framework:

$$K_S \exists x[(C_A x \wedge P) \vee (\neg C_A x \wedge P)]^+ \vDash K_S \Diamond \exists x(C_A x \wedge Px) \wedge K_S \Diamond \exists x(\neg C_A x \wedge Px)$$

Proof.

According to Fact 3., we infer from $K_S \exists x[(C_A x \wedge P) \vee (\neg C_A x \wedge P)]^+$ to conclude that $K_S \exists x[(C_A x \wedge P) \vee (\neg C_A x \wedge P)]^+$
Suppose we have a model $\mathcal{M}$, a state s and R which is an equivalence relation defined on W.

And then assume that $\mathcal{M}, s \vDash K_S \exists x[(C_A x \wedge P) \vee (\neg C_A x \wedge P)]^+$. It follows that $\forall i \in s$, $\mathcal{M}, s_i^K \vDash \exists x[(C_A x \wedge P) \vee (\neg C_A x \wedge P)]^+$. And then we can conclude that there are non-empty states t and t' s.t. $t \cup t' = s_i^K[x/h]$ for some h and for $\forall i \in s$, $\mathcal{M}, t \vDash C_A x \wedge Px$ and $\mathcal{M}, t' \vDash \neg C_A x \wedge Px$.

Since t is a non-empty subset of $s_i^K[x/h]$, we directly get $s_i^K[x/h] \vDash \Diamond(C_A x \wedge Px)$ for some h and for every $i \in s$. It follows that $\mathcal{M}, s_i^K \vDash \exists x \Diamond (C_A x \wedge Px)$. Finally since for $\forall i \in s$ the knowledge state s_i^K supports $\exists x \Diamond (C_A x \wedge Px)$, we can get $\mathcal{M}, s \vDash K_S \exists x \Diamond (C_A x \wedge Px)$.

It is similar to prove $M, s \vDash K \exists x \Diamond (\neg C_A x \wedge Px)$. □

We need to highlight the effect of knowledge state in this example. First, the knowledge state s_i^K in this example represents part of Susan's knowledge that learned by Tom, namely the state of Susan's knowledge based on the speaker's information state. By the framework QBSEL, the ignorance inference in (22) is due to the assumption of neglecting zero rather than the conversation or reasoning about something else.

6 Conclusion

Monotonic reasoning under attitude verbs sometimes appears to fail, yet sometimes doesn't. Especially for the deontic and bouletic verbs, there are plenty of arguments about whether these verbs block the monotonic reasoning we have reviewed in Sec.2. This brought people to consider further whether the classical monotonic semantics for these verbs is correct. Therefore the issue about monotonicity becomes important to determine whether the classical monotonic semantics should be accepted.

In this paper, we adopt a strategy that takes us away from the debate, and examine whether similar phenomena exist in other domains. We hence look at the epistemic domain, where the classical semantics is less problematic in its predictions about monotonicity and is normally accepted. A puzzle is presented to show that the monotonic reasoning under the verb 'know' still can lead to wrong decisions. Therefore we have reasons to believe the problem about monotonicity is a matter of pragmatics, rather than semantics. We consider the oddness is from some additional inferences licensed by the conclusion of monotonicity pragmatically. These inferences give rise to some pragmatic effect which influence our prediction or decision in practice. However, they are not always bad. We also present examples to show that some correct predictions can be deduced from monotonicity. In order to explain how the pragmatic effect is generated, we first propose an analysis that the predicate can convey a disjunctive meaning which could trigger some modal inferences, e.g., the ignorance inference. Then we develop a framework QBSEL to provide a logic-based account for capturing such inferences.

In epistemic setting, it is uncontroversial that these failure are apparent and just a matter of pragmatics. But then once a pragmatic story has been developed can also be applied to other domain. In doing so, we expect to defend the classically monotonic semantics. We believe it is not so bad, at least for the problem w.r.t monotonicity.

References

1. Maria Aloni. Logic and conversation: the case of free choice. Presented at LINGUAE Seminar (Paris, 2019), Principles of Formal Semantics (Stockholm, 2019), Roots of Pragmasemantics (Szklarska Poreba, 2020), Logic Seminar (Helsinki, 2020), ICLA (Chennai, 2021), LoDE (Utrecht 2021)., 2021.
2. Maria Aloni and Peter van Ormondt. Modified numerals and split disjunction: the first-order case. Manuscript, ILLC, University of Amsterdam, 2021.
3. Maria Aloni and Robert Van Rooij. Free choice items and alternatives. In Proceedings of the KNAW academy colloquium: Cognitive foundations of interpretation, pages 5–26. Edita KNAW Amsterdam, 2007.
4. Nicholas Asher. A typology for attitude verbs and their anaphoric properties. Linguistics and Philosophy, 10(2):125–197, 1987.
5. Daniel Büring. The least at least can do. In West Coast Conference on Formal Linguistics (WCCFL), volume 26, pages 114–120. Citeseer, 2008.
6. Fabrizio Cariani. ?ought? and resolution semantics. Noûs, 47(3):534–558, 2013.
7. Pual Dekker. Transentential Meditations: Ups and Downs in Dynamic Semantics. PhD thesis, ILLC, University of Amsterdam, 1993.
8. Ronald Fagin and Joseph Y Halpern. Belief, awareness, and limited reasoning. Artificial intelligence, 34(1):39–76, 1987.

9. Gerald Gazdar. Pragmatics, implicature, presuposition and lógical form. Critica, 12(35):113–122, 1979.
10. Herbert P Grice. Logic and conversation. In Speech acts, pages 41–58. Brill, 1975.
11. Irene Heim. Presupposition projection and the semantics of attitude verbs. Journal of semantics, 9(3):183–221, 1992.
12. Jaakko Hintikka. Semantics for propositional attitudes. In Models for modalities, pages 87–111. Springer, 1969.
13. Mark Jago. The impossible: An essay on hyperintensionality. OUP Oxford, 2014.
14. Hans Kamp. Iv—free choice permission. Proceedings of the Aristotelian Society, 74(1):57–74, 1974.
15. Daniel Lassiter. Measurement and modality: The scalar basis of modal semantics. Ph. D. thesis, New York University, 2011.
16. Dmitry Levinson. Probabilistic model-theoretic semantics for'want'. In Semantics and Linguistic Theory, volume 13, pages 222–239, 2003.
17. Bill MacCartney. Natural language inference. Stanford University, 2009.
18. Jon Ander Mendia. Conveying ignorance: Ignorance inferences with superlative numeral modifiers. Proceedings of ConSOLE XXIII, 150:174, 2015.
19. Léo Picat and Salvador Mascarenhas. On the interplay between interpretation and reasoning in compelling fallacies. 2020.
20. Alf Ross. Imperatives and logic. Journal of Symbolic Logic, 6(3):105–106, 1941.
21. Uli Sauerland. Scalar implicatures in complex sentences. Linguistics and Philosophy, 27(3):367–391, 2004.
22. Zhiqiang Sun and Fenrong Liu. The inference pattern mou in mohist logic: A monotonicity reasoning view. Roczniki Filozoficzne, 68(4):257–270, 2020.
23. Jouko Väänänen. Dependence logic: A new approach to independence friendly logic, volume 70. Cambridge University Press, 2007.
24. Johan van Benthem. A brief history of natural logic.
25. Johan van Benthem and Fenrong Liu. New logical perspectives on monotonicity. In Tsinghua Interdisciplant Workshop on Logic, Language, and Meaning: Monotonicity in Logic and Language, pages 1–12. Springer, 2020.
26. Jan Van Eijck. Syllogistics= monotonicity+ symmetry+ existential import. preprint May, 2005.
27. Kai Von Fintel. Npi licensing, strawson entailment, and context dependency. Journal of semantics, 16(2):97–148, 1999.
28. Fan Yang and Jouko Väänänen. Propositional team logics. Annals of Pure and Applied Logic, 168(7):1406–1441, 2017.
29. Thomas Ede Zimmermann. Free choice disjunction and epistemic possibility. Natural language semantics, 8(4):255–290, 2000.