

## 1 Categorical Well-formedness: \*NT

Let us assume a successor model with features and the Boolean semiring.

Name	$S$	$\oplus$	$\otimes$	0	1
Boolean	{ <b>true</b> , <b>false</b> }	$\vee$	$\wedge$	<b>false</b>	<b>true</b>

Consider the constraint \*NT defined below.

$$*NT \stackrel{\text{def}}{=} \exists x, y [x \triangleleft y \wedge \text{nasal}(x) \wedge \text{cons}(y) \wedge \text{voiced}(y)] \quad (1)$$

“There exist  $x, y$  such that  $y$  succeeds  $x$  and  $x$  is a nasal and  $y$  is a consonant and  $y$  is voiced.”

“Voiceless consonants after nasals are marked.”

### 1.1 Example evaluation: endond

$$\llbracket *NT \rrbracket(\text{endond}) = \bigvee_{x \in D, y \in D} [x \triangleleft y \wedge \text{nasal}(x) \wedge \text{cons}(y) \wedge \text{voiced}(y)]$$

In weighted logic, existential quantification is interpreted as semiring addition, which is disjunction in the Boolean semiring. Hence the big disjunction symbol which conjoins *every* pair  $(x, y) \in D \times D$ . The disjunction of these terms is only **true** if at least one disjunct is **true**. Only when every disjunct is **false** is the entire term **false**.

The conjunction inside the outermost disjunction is only **true** when  $x = 2, y = 3$  and  $x = 5, y = 6$ . In every other case, it is **false**. Thus, the disjuncts contain two **true**s with every other disjunct being **false**. It follows that their disjunction is **true** and so  $\llbracket *NT \rrbracket(\text{endond}) = \text{true}$ . In other words, *endond* does violate this constraint.

### 1.2 Example evaluation: enton

$$\llbracket *NT \rrbracket(\text{enton}) = \bigvee_{x \in D, y \in D} [x \triangleleft y \wedge \text{nasal}(x) \wedge \text{cons}(y) \wedge \text{voiced}(y)]$$

The calculation proceeds as above. In the model of this word, the interior conjunction is **false** for every pair  $x, y$ . It follows that their disjunction is **false** and so  $\llbracket *NT \rrbracket(\text{enton}) = \text{false}$ . In other words, *enton* does not violate \*NT.

## 2 Counting violations: \*NT

Let us assume a successor model with features and the Natural semiring.

Name	$S$	$\oplus$	$\otimes$	0	1
Natural	$\mathbb{N}$	$+$	$\times$	0	1

We consider now the constraint \*NT shown below.

$$*NT \stackrel{\text{def}}{=} \exists x, y [x \triangleleft y \wedge \text{nasal}(x) \wedge \text{cons}(y) \wedge \text{voiced}(y)] \quad (2)$$

This looks just like Equation 1! In the Natural semiring, existential quantification is interpreted as *addition*, and conjunction is interpreted as *multiplication*.

## 2.1 Example evaluation: endond

$$\llbracket *NT \rrbracket(\text{endond}) = \sum_{x \in D, y \in D} \left[ x \triangleleft y \times \text{nasal}(x) \times \text{cons}(y) \times \text{voiced}(y) \right]$$

Recall that in weighted logic, if the model satisfies an atomic predicate under some assignment, it takes on the semiring unit value, which in the Natural semiring is 1. Thus, when  $x = 2, y = 3$  we have the following computation:

$$\begin{aligned} & x \triangleleft y \times \text{nasal}(x) \times \text{cons}(y) \times \text{voiced}(y) \\ &= 2 \triangleleft 3 \times \text{nasal}(2) \times \text{cons}(3) \times \text{voiced}(3) \\ &= 1 \times 1 \times 1 \times 1 \\ &= 1 \end{aligned}$$

On the other hand, when when  $x = 2, y = 4$  we have the following computation:

$$\begin{aligned} & x \triangleleft y \times \text{nasal}(x) \times \text{cons}(y) \times \text{voiced}(y) \\ &= 2 \triangleleft 4 \times \text{nasal}(2) \times \text{cons}(4) \times \text{voiced}(4) \\ &= 0 \times 1 \times 0 \times 1 \\ &= 0 \end{aligned}$$

It follows that the terms over all  $x, y$  pairs evaluate to zero except when  $x = 2, y = 3$  and  $x = 5, y = 6$ . Those terms will evaluate to 1. The sum thus is 2, and we have  $\llbracket *NT \rrbracket(\text{endond}) = 2$ . In other words, *endond* violates *\*NT* twice.

## 2.2 Example evaluation: enton

$$\llbracket *NT \rrbracket(\text{entom}) = \sum_{x \in D, y \in D} \left[ x \triangleleft y \times \text{nasal}(x) \times \text{cons}(y) \times \text{voiced}(y) \right]$$

The calculation proceeds as above. No term in the sum evaluates to 1. It follows that  $\llbracket *NT \rrbracket(\text{entom}) = 0$ . In other words, *entom* violates *\*NT* zero times.

## 3 Gradient well-formedness: \*NT

Let us assume a successor model with features and the Real Interval semiring.

Name	$S$	$\oplus$	$\otimes$	0	1
Real Interval	$[0, 1]$	+	$\times$	0	1

We define some user defined predicates to make the final equation more readable.

$$\text{NT}(x, y) \stackrel{\text{def}}{=} x \triangleleft y \wedge \text{nasal}(x) \wedge \text{cons}(y) \wedge \neg \text{voiced}(y) \quad (3)$$

$$\text{adj}\neg\text{NT}(x, y) \stackrel{\text{def}}{=} x \triangleleft y \wedge (\neg \text{nasal}(x) \vee \neg \text{cons}(y) \vee \text{voiced}(y)) \quad (4)$$

$$\text{nonadj}(x, y) \stackrel{\text{def}}{=} \neg(x \triangleleft y) \quad (5)$$

Observe these are *mutually exclusive* conditions – any  $(x, y) \in D \times D$  satisfies exactly one of Equations 3, 4, and 5. We now define the constraint *\*NT* with the Real Interval semiring as follows.

$$*NT \stackrel{\text{def}}{=} \forall x, y \left[ (\text{NT}(x, y) \wedge 0.5) \vee (\text{adj}\neg\text{NT}(x, y)) \vee (\text{nonadj}(x, y)) \right] \quad (6)$$

### 3.1 Example evaluation: endond

$$\llbracket^* \text{NT} \rrbracket(\text{endond}) = \prod_{x \in D, y \in D} \left[ (\text{NT}(x, y) \times 0.5) + (\text{adj-NT}(x, y)) + (\text{nonadj}(x, y)) \right]$$

Ultimately this is a product of sums. The terms in each sum are mutually exclusive. In other words, for any  $x, y$  exactly one of  $(\text{NT}(x, y) \times 0.5)$ ,  $(\text{adj-NT}(x, y))$ ,  $(\text{nonadj}(x, y))$  is nonzero. Thus exactly one of those terms determines the sum since the other two are zero.

By definition of the semantics of weighted logic, if  $x, y$  satisfy  $(\text{adj-NT}(x, y))$  or  $(\text{nonadj}(x, y))$  then those will evaluate to 1 and the sum of these three terms will also be 1. On the other hand, if  $x, y$  satisfies  $\text{NT}(x, y)$  then the first term will evaluate to  $1 \times 0.5 = 0.5$ .

Finally, we must take the product of these sums. Exactly two terms in the product are 0.5; the others are 1. Thus  $\llbracket^* \text{NT} \rrbracket(\text{endond}) = 0.25$ . In other words, each occurrence of ND reduces the well-formedness by one-half.

### 3.2 Example evaluation: enton

The computation is the same as above, but in this case, every term in the product will be 1 and therefore this word will have a gradient well-formedness score of 1.