

Комплекс Д32, Анна Полянская

1) Предложите список отношений для русского языка и приведите хотя бы один пример на каждое отношение.

первый уровень аннотации

1. равенство (эквивалент) / перифраз
 - а. собака ест – пес кушает
2. прямое следствие
 - а. собака на самоизоляции – собака сидит дома
3. обратное следствие
 - а. у собаки была операция – собака отходит от наркоза
4. противоречие
 - а. собака сидит дома на карантине – собака гуляет в парке
5. не связано (нейтрально)
 - а. в доме живет кошка – в доме нет мышей

второй уровень аннотации:

1. Кореференция
 - а. Нет: собака грызет кость – кошка поймала мышь
 - б. Есть: собака грызет кошку – кошка поймала мышь
2. Belief (принятое следствие)
 - а. Нет: собака хорошо поела – собака легла спать
 - б. Есть: собака много играла днем, потом легла спать – собака устала

2) обоснуйте свой выбор отношений.

Я предполагаю, что такой набор отношений должен повысить уровень согласия, а наличие второго уровня аннотации поможет объяснить причины разногласий, выявить явные ошибки разметки, и, возможно, дадут какое-то понимание культурного кода откуда они берутся (потому что для разных людей одни и те же ситуации могут значить разные вещи. Минимальный пример, между выражениями «министр культуры не встала во время государственного гимна» и «министр культуры выразила свою политическую позицию» в одной культуре будет прямая связь, в другой – belief, в третьей – нет связи)

Вероятно, еще стоит разделять прямое следствие, обратное следствие и пресуппозицию (например, собака отходит от наркоза – собака существует/собака была под наркозом), потому что последнее выделяет просто какие-то куски из определений составляющих высказывания, озвучивание пресуппозиций не особо распространено в речи, зато может дать много информации о т.н. языковой картине мира (по крайней мере для данного языка).

3) предположите (и обоснуйте своё предположение), какие из предложенных Вами отношений будут вызывать наибольшие проблемы у людей-аннотаторов, а какие у SNLI/MultiNLI

1. У SNLI/MultiNLI могут возникать сложности с belief, потому что автоматически можно определить далеко не все стереотипы, заложенные в языке, и равенством и кореференцией, потому что в языке может быть очень много синонимов и способов перифразы.
2. У людей скорее всего будет больше всего проблем в нейтральном случае, как раз по причине наличия кореференции и belief, которые будут вызывать ощущение следствия или противоречия.