

Measuring Expressions of Uncertainty in Radiology Texts for Natural Language Processing Applications

BE Chapman¹, A Gentili², JY Chen²,
A Miyakoshi², WW Chapman¹

1. University of Utah, Salt Lake City, UT
2. UC, San Diego and San Diego VA, San Diego, CA



Background

- Natural language processing (NLP) is
 - an important tool for extracting structured information from radiology texts
- pyConTextNLP uses linguistic cues to determine whether a finding is negated, asserted, or uncertain



131	highly suggestive	DEFINITE_EXISTENCE
132	obvious	DEFINITE_EXISTENCE
133	positive study for	DEFINITE_EXISTENCE
134	represents	DEFINITE_EXISTENCE
135	believe	DEFINITE_EXISTENCE
136	can see	DEFINITE_EXISTENCE
137	gross evidence	DEFINITE_EXISTENCE
138	high probability	DEFINITE_EXISTENCE
139	is positive	DEFINITE_EXISTENCE
140	reveals	DEFINITE_EXISTENCE
141	are ruled out	DEFINITE_NEGATED_EXISTENCE
142	can be ruled out	DEFINITE_NEGATED_EXISTENCE
143	could be ruled out	DEFINITE_NEGATED_EXISTENCE
144	free	DEFINITE_NEGATED_EXISTENCE
145	has been ruled out	DEFINITE_NEGATED_EXISTENCE
146	have been ruled out	DEFINITE_NEGATED_EXISTENCE
147	is negative	DEFINITE_NEGATED_EXISTENCE
148	is ruled out	DEFINITE_NEGATED_EXISTENCE



Challenge

Compiling linguistic cues that accurately represent spectrum of uncertainty

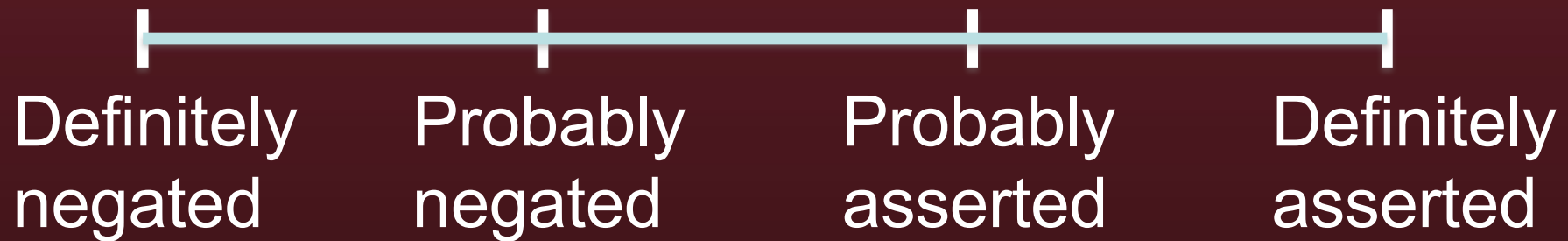


Definitely
negated

Definitely
asserted

Objective

Compare probabilities assigned by radiologists against four categories defined in pyConTextNLP



Evaluation

- 133 pyConTextNLP cues
- 108 cues translated from Swedish clinical texts



Evaluation

- Three radiologists separately reviewed cues presented in random order
 - Assigned single-point probabilities
 - Assigned probability ranges
 - blinded to single-point responses



Existence Cue Point Mapping

241 of 242 cues remaining to be mapped; Annotator: brian

Please assign the probability (0-100) a subject would have a finding, given this cue modifier.

If the term seems nonsensical or if you have other concerns, please enter a comment. There is no need to provide a comment otherwise.

cue value: **"nothing"**

Probability:	<input type="text" value=""/>
Comment:	<input type="text"/>
<input type="button" value="Save and Next unmarked"/>	
CUEID:	<input type="text" value="156"/>

Existence Cue Range Mapping

240 of 242 cues remaining to be mapped; Annotator: brian

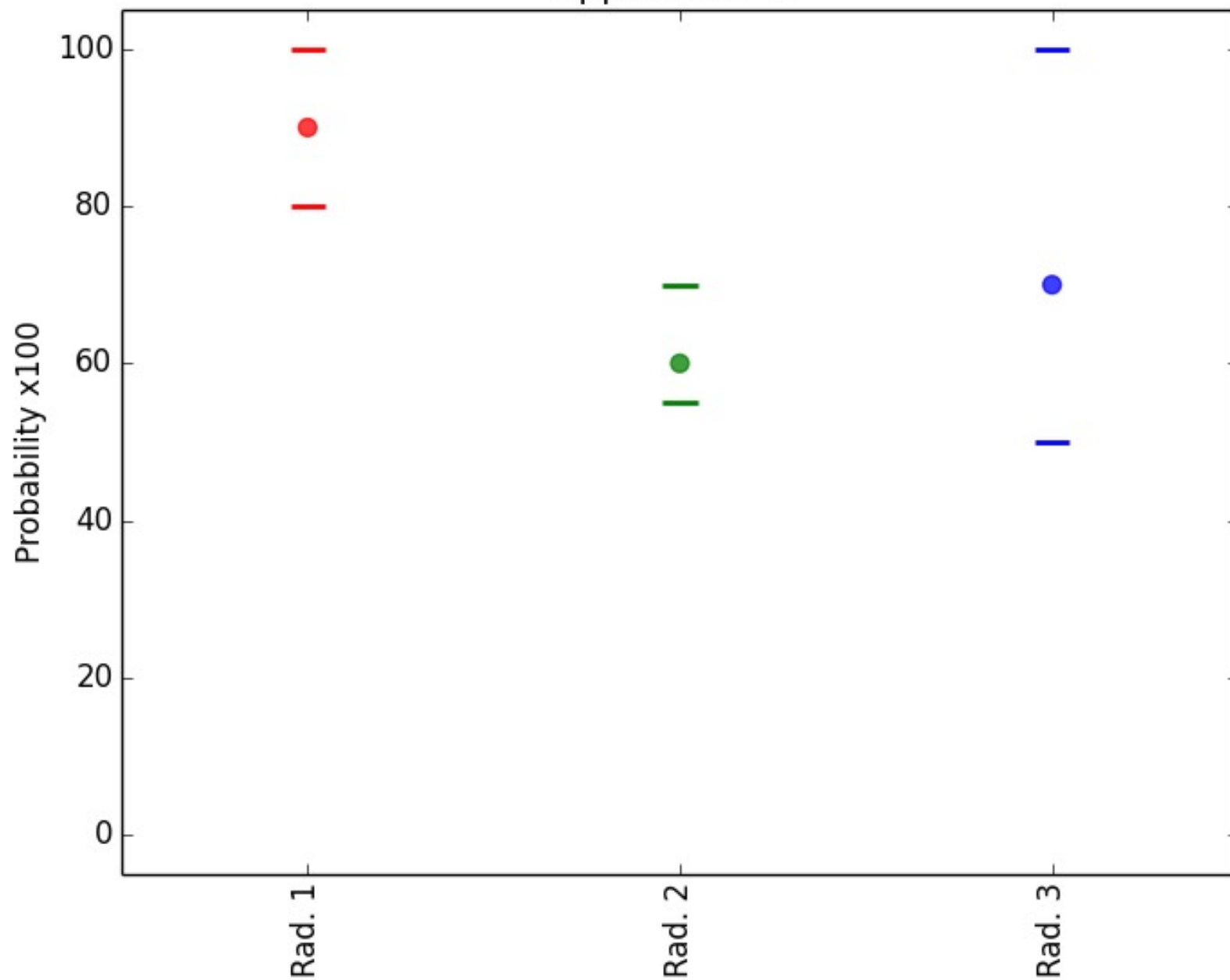
Please assign the probability (0-100) a subject would have a finding, given this cue as a modifier.

If the term seems nonsensical or if you have other concerns, please enter a comment. There is no need to provide a comment otherwise.

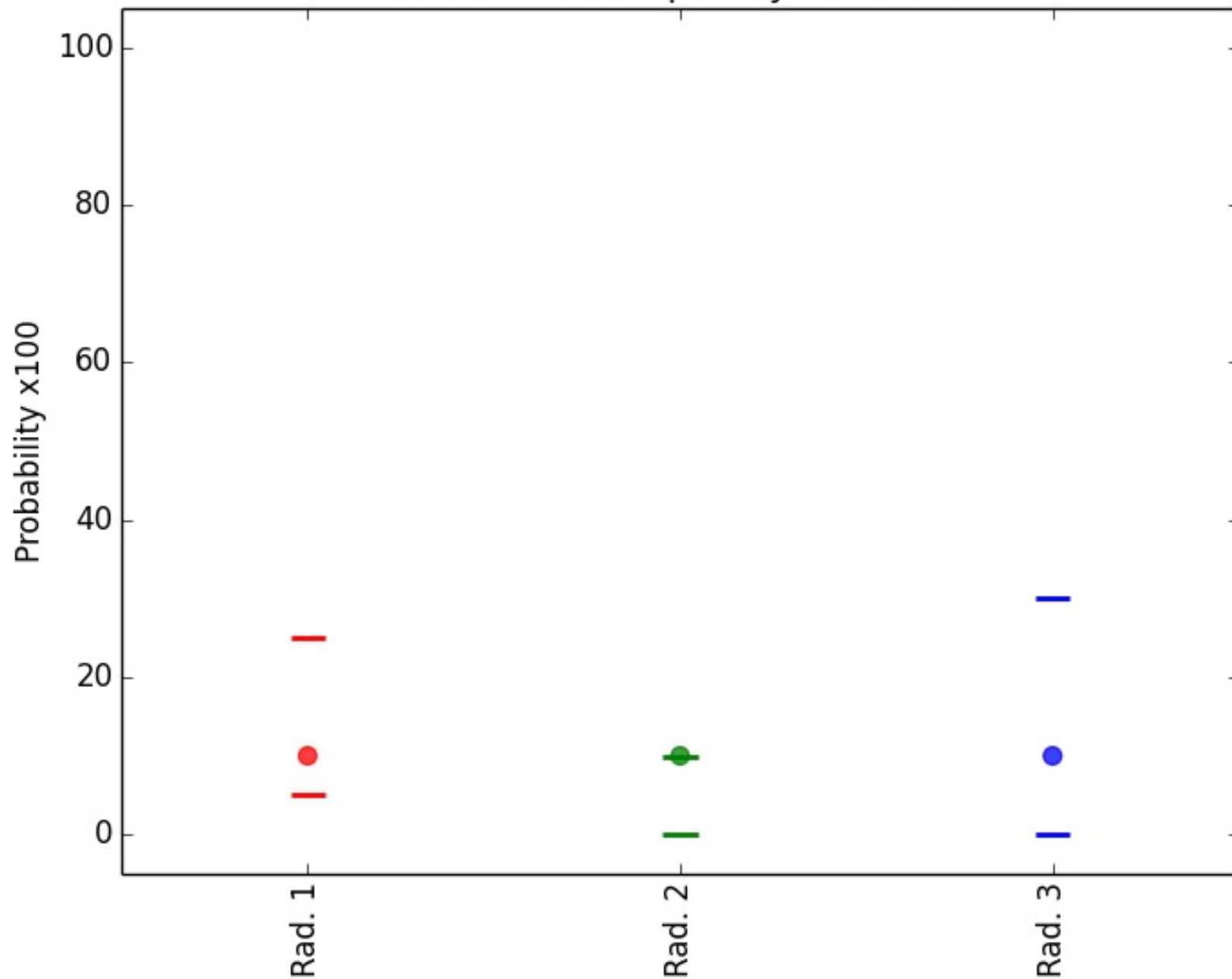
Cue value: "may be"

LowProbability:	<input type="text"/>
HighProbability:	<input type="text"/>
Comment:	<input type="text"/>
<input type="button" value="Save and Next unmarked"/>	

appears to be



cannot be completely excluded



Results

Pairwise comparisons of single-point probabilities

- Very similar mean values
 - mean difference of 0.012
- Large variability in standard deviation
 - mean standard deviation of 0.21

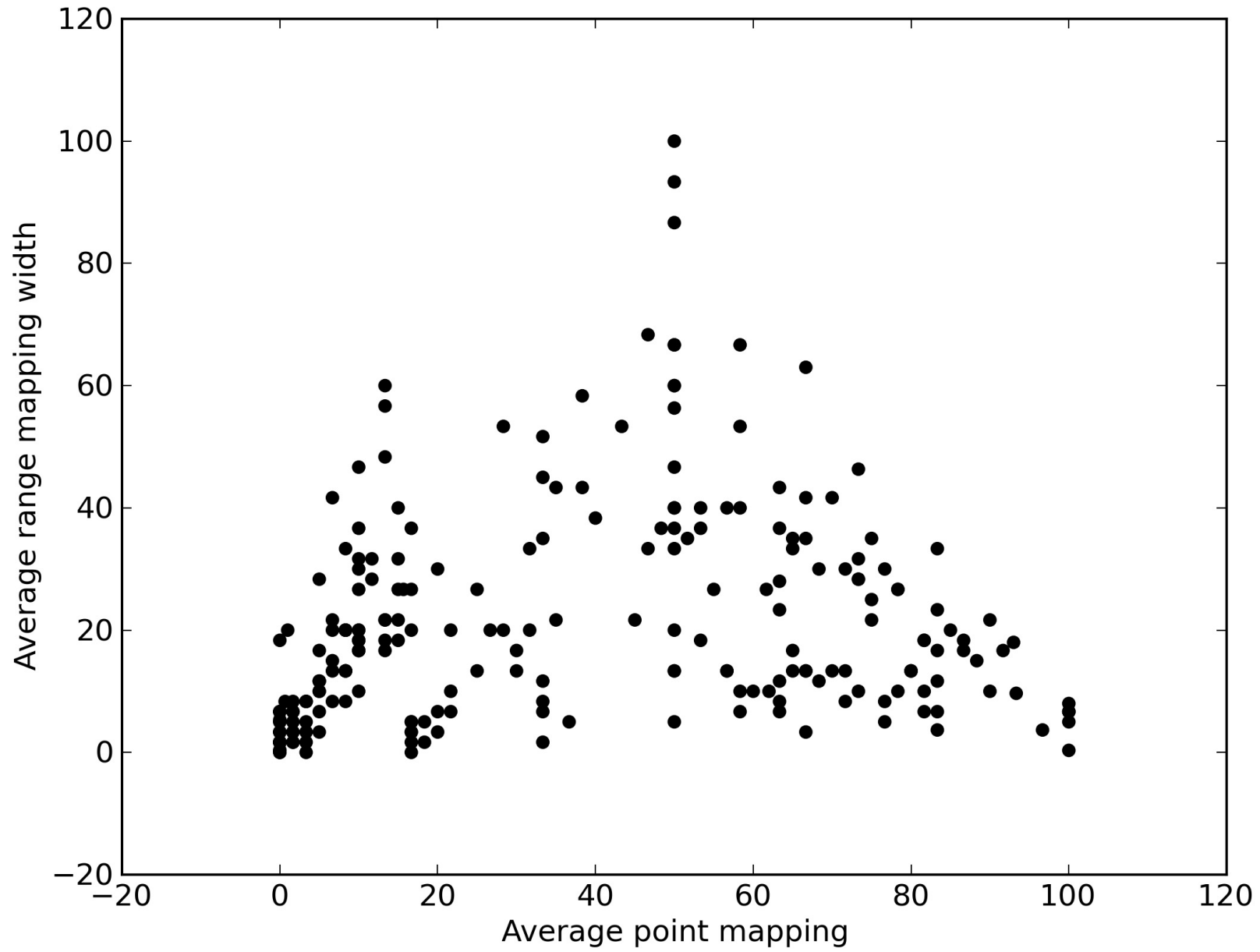


Results

Range mappings

- Small differences with center of ranges and point mappings
 - mean location (-0.0035)
- Large variability in widths
 - mean standard deviation of 0.21





Results

Point probabilities by categories
mean (std):

- definitely negated 0.078 (0.11)
- probably negated 0.17 (0.16)
- probably asserted 0.71 (0.11)
- definitely asserted 0.91 (0.083).



Conclusions

- Radiologist had high consistency in point probabilities
 - but variance was high
- Overlap occurred between definitely and probably negated categories
- Our model of uncertainty could be improved by adding a fifth category of “ambivalent”
 - Capture highly uncertain existence cues with probabilities near 0.5

