Computational Approaches Towards Monotonicity in Sentence-Picture Verification

Jakub Szymanik

Joint work with Marcin Zajenkowski
TLS 13

university of groningen

Outline

Outline

university of groningen

NL determiners

1. All poets have low self-esteem.
2. Some dean danced nude on the table.
3. At least 3 grad students prepared presentations.
4. An even number of the students saw a ghost.
5. Most of the students think they are smart.
6. Less than half of the students received good marks.

Monotone quantifiers

Definition

Q is increasing iff, for any $A \subseteq A'$, $\mathsf{Q}(A, B)$ entails $\mathsf{Q}(A', B)$.

Monotone quantifiers

Definition

Q is increasing iff, for any $A \subseteq A'$, $\mathsf{Q}(A, B)$ entails $\mathsf{Q}(A', B)$.

Example

1. Every child is dirty.
2. Every boy is dirty.

Monotone quantifiers

Definition

Q is increasing iff, for any $A \subseteq A'$, $\mathsf{Q}(A, B)$ entails $\mathsf{Q}(A', B)$.

Example

1. Every child is dirty.
2. Every boy is dirty.

Definition

Q is decreasing iff, for any $A \subseteq A'$, $\mathsf{Q}(A', B)$ entails $\mathsf{Q}(A, B)$.

Monotone quantifiers

Definition

Q is increasing iff, for any $A \subseteq A'$, $\mathsf{Q}(A, B)$ entails $\mathsf{Q}(A', B)$.

Example

1. Every child is dirty.
2. Every boy is dirty.

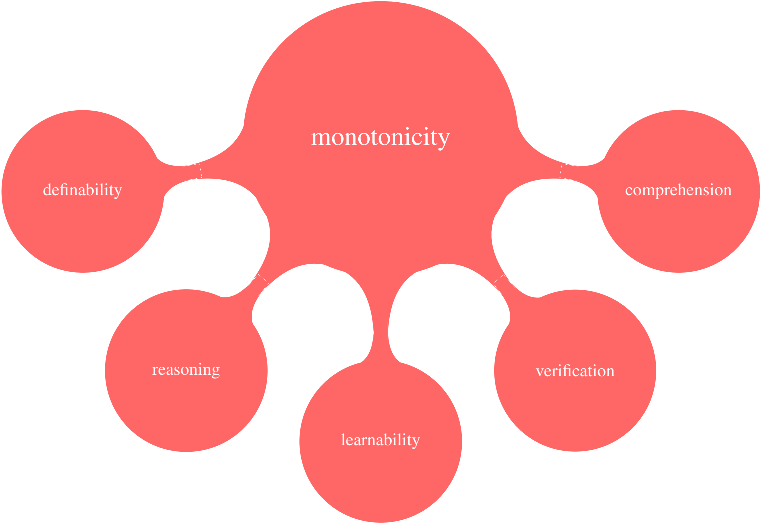
Definition

Q is decreasing iff, for any $A \subseteq A'$, $\mathsf{Q}(A', B)$ entails $\mathsf{Q}(A, B)$.

Example

1. No child is dirty.
2. No boy is dirty.

Monotonicity – key property in logic and language



Barwise and Cooper's observation

Observation

Monotonicity relates to verification procedures.

Barwise and Cooper's observation

Observation

Monotonicity relates to verification procedures.

Example (witness set)

1. Some cars are green.
2. No cars are red.

Barwise and Cooper's observation

Observation

Monotonicity relates to verification procedures.

Example (witness set)

1. Some cars are green.
2. No cars are red.

Hypothesis

'Response latencies for verification tasks involving decreasing quantifiers would be somewhat greater than for increasing quantifiers' (p. 192).

Barwise and Cooper, Generalized Quantifiers and Natural Language, Linguistics and Philosophy, 1981

Truth-values??

Question

What about truth-value?

Outline

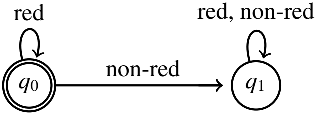
Outline

Simplicity

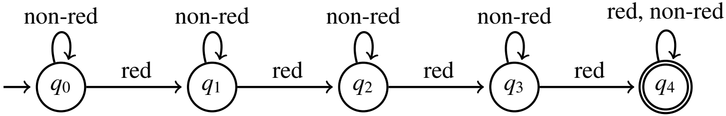
- Simple quantifiers can be computed by simple automata.
- Encoding natural counting strategies.
- We restrict ourselves to precise counting.

Van Benthem, Essays in logical semantics, 1986

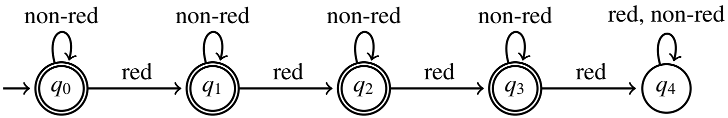
Every dot is red.



More than 3 dots are red.



Fewer than 4 dots are red.

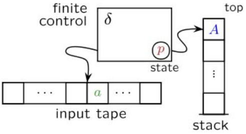


Proportional quantifiers

1. More than half of the dots are red.
2. Fewer than half of the dots are red.

Proportional quantifiers

1. More than half of the dots are red.
2. Fewer than half of the dots are red.



- Not computable by finite-automata.
- We need working memory.
- Simple push-down automata will do.

Psychological plausibility

Observation

The more complex automata the longer RT and the greater WM involvement.

McMillan et al., Neural basis for generalized quantifiers comprehension, Neuropsychologia, 2005

Szymaniki & Zajenkowski, Comprehension of simple quantifiers. Empirical evaluation of a computational model, Cognitive Science, 2010

Zajenkowski et al., A computational approach to quantifiers as an explanation for some language impairments in schizophrenia, JCD, 2011

Outline

Design

- 4 different quantifiers:

Design

- 4 different quantifiers:
  - Cardinal quantifiers (“more than 7”,“fewer than 8”);

Design

- 4 different quantifiers:
  - Cardinal quantifiers (“more than 7”,“fewer than 8”);
  - Proportional (“more than half”, “fewer than half”).

Design

- 4 different quantifiers:
  - Cardinal quantifiers (“more than 7”,“fewer than 8”);
  - Proportional (“more than half”, “fewer than half”).
- Upward monotone vs. downward monotone.

Design

- 4 different quantifiers:
  - Cardinal quantifiers ("more than 7","fewer than 8");
  - Proportional ("more than half", "fewer than half").
- Upward monotone vs. downward monotone.
- True vs. false.

Design

- 4 different quantifiers:
  - Cardinal quantifiers (“more than 7”,“fewer than 8”);
  - Proportional (“more than half”, “fewer than half”).
- Upward monotone vs. downward monotone.
- True vs. false.
- Subjects were timed when asked to decide if true.

Design

- 4 different quantifiers:
  - Cardinal quantifiers (“more than 7”,“fewer than 8”);
  - Proportional (“more than half”, “fewer than half”).
- Upward monotone vs. downward monotone.
- True vs. false.
- Subjects were timed when asked to decide if true.
- Reading and verification time.

Predictions

1. RT increase along with the complexity.

Predictions

1. RT increase along with the complexity.
2. Complexity influenced by (monotonicity $\times$ truth-value):
   - In the case of the cardinal sentences,
   - but not the proportional sentences.

Predictions in details

1. "More than 7": true > false (8>7).

Predictions in details

1. "More than 7": true > false (8>7).
2. "Fewer than 8": true < false (7<8).

Predictions in details

1. “More than 7”: true > false (8>7).
2. “Fewer than 8”: true < false (7<8).
3. No difference between proportional quantifiers.

Predictions in details

1. “More than 7”: true > false (8>7).
2. “Fewer than 8”: true < false (7<8).
3. No difference between proportional quantifiers.
4. Proportional quantifiers > cardinal quantifiers.

Outline

university of groningen

- 61 native Polish-speaking adults (33 female).
- Volunteers: undergraduates from the University of Warsaw.
- The mean age: 21.42 years (SD = 3.22).
- Each participant tested individually.

Materials

16 grammatically simple propositions in Polish, like:

1. More than 7 cars are blue.
2. Fewer than 8 cars are yellow.
3. More than half of the cars are red.
4. Fewer than half of the cars are black.

Materials continued

More than half of the cars are yellow.

An example of a stimulus used in the first study

Procedure

- Each quantifier was presented in 10 trials.

Procedure

- Each quantifier was presented in 10 trials.
- The sentence true in the picture in half of the trials.

Procedure

- Each quantifier was presented in 10 trials.
- The sentence true in the picture in half of the trials.
- Quantity of target items near the criterion of validation.

Procedure

- Each quantifier was presented in 10 trials.
- The sentence true in the picture in half of the trials.
- Quantity of target items near the criterion of validation.
- Practice session followed by the experimental session.

Procedure

- Each quantifier was presented in 10 trials.
- The sentence true in the picture in half of the trials.
- Quantity of target items near the criterion of validation.
- Practice session followed by the experimental session.
- Each quantifier problem was given one 15.5 s event.

Procedure

- Each quantifier was presented in 10 trials.
- The sentence true in the picture in half of the trials.
- Quantity of target items near the criterion of validation.
- Practice session followed by the experimental session.
- Each quantifier problem was given one 15.5 s event.
- Subjects were asked to decide the truth-value.

Procedure

- Each quantifier was presented in 10 trials.
- The sentence true in the picture in half of the trials.
- Quantity of target items near the criterion of validation.
- Practice session followed by the experimental session.
- Each quantifier problem was given one 15.5 s event.
- Subjects were asked to decide the truth-value.
- Reading and verification stages.

Procedure

- Each quantifier was presented in 10 trials.
- The sentence true in the picture in half of the trials.
- Quantity of target items near the criterion of validation.
- Practice session followed by the experimental session.
- Each quantifier problem was given one 15.5 s event.
- Subjects were asked to decide the truth-value.
- Reading and verification stages.
- Cardinal and proportional quantifiers were equivalent.

Outline

university of groningen

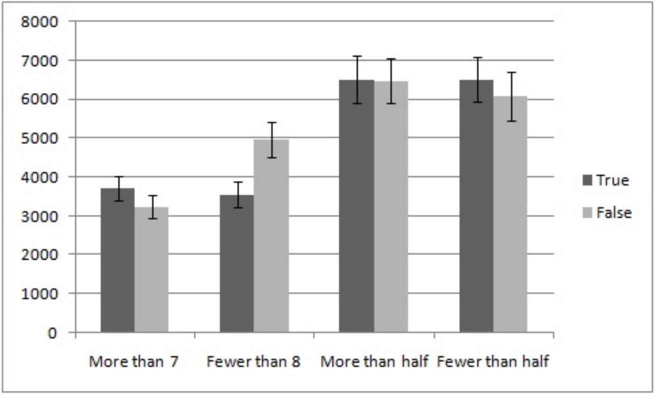
Reading times were similar

Quantifier	M	SD
More than 7	3430	1497
Fewer than 8	3577	1793
More than half	3695	1698
Fewer than half	3908	1639

Verification times

Quantifier		M	SD
More than 7	True	3710	1230
	False	3241	1167
	Overall	3475	955
Fewer than 8	True	3564	1311
	False	4971	1843
	Overall	4259	1312
More than half	True	6511	2454
	False	6475	2195
	Overall	6493	1959
Fewer than half	True	6509	2378
	False	6084	2464
	Overall	6296	2090

Interaction effect



Outline

False/True interaction

Refinement on Barwise & Cooper.

False/True interaction

Refinement on Barwise & Cooper.

- $RT(\exists_T^{>7}) > RT(\exists_F^{>7})$

False/True interaction

Refinement on Barwise & Cooper.

- $RT(\exists_T^{>7}) > RT(\exists_F^{>7})$

False/True interaction

Refinement on Barwise & Cooper.

- $RT(\exists_T^{>7}) > RT(\exists_F^{>7})$ 👍
- $RT(\exists_T^{<8}) < RT(\exists_F^{<8})$.

False/True interaction

Refinement on Barwise & Cooper.

- $RT(\exists_T^{>7}) > RT(\exists_F^{>7})$ 👍
- $RT(\exists_T^{<8}) < RT(\exists_F^{<8})$. 👍

False/True interaction

Refinement on Barwise & Cooper.

- $RT(\exists_T^{>7}) > RT(\exists_F^{>7})$ 👍
- $RT(\exists_T^{<8}) < RT(\exists_F^{<8})$. 👍
- No difference in proportional quantifiers.

False/True interaction

Refinement on Barwise & Cooper.

- $RT(\exists_T^{>7}) > RT(\exists_F^{>7})$ 👍
- $RT(\exists_T^{<8}) < RT(\exists_F^{<8})$. 👍
- No difference in proportional quantifiers. 👍

False/True interaction

Refinement on Barwise & Cooper.

- $RT(\exists_T^{>7}) > RT(\exists_F^{>7})$
- $RT(\exists_T^{<8}) < RT(\exists_F^{<8})$.
- No difference in proportional quantifiers.
- Complexity hypothesis.

False/True interaction

Refinement on Barwise & Cooper.

- $RT(\exists_T^{>7}) > RT(\exists_F^{>7})$
- $RT(\exists_T^{<8}) < RT(\exists_F^{<8})$.
- No difference in proportional quantifiers.
- Complexity hypothesis.

Take home message

- Complexity influences verification
- Monotonicity $\times$ truth-value influences complexity

Thanks!

Outline

Just and Carpenter 1971

Observation

Processing time of negative quantifiers is greater than processing time of affirmative quantifiers.

Just & Carpenter, Comprehension of negation with quantification, Journal of Verbal Learning and Verbal Behavior, 1971

3 kinds of sentences

1. Syntactic negatives with particle:
   - The dots are red.
   - The dots aren’t red.
2. Syntactic negatives without particle:
   - Many of the dots are red.
   - Few of the dots are red.
3. Semantic negatives:
   - A majority of the dots are red.
   - A minority of the dots are red.

Only some pairs contrasted w.r.t. monotonicity:

1. All of the dots are red.
2. None of the dots are red.

Most of the material was based on negativity vs. affirmativity.

Negativity is Marked, not only linguistically

Example

How tall are you? but not How short are you?

Example (Squirrel Monkeys)

1. If everything is black, choose the biggest object.
2. If everything is white, choose the smallest object.

Once trained, monkey were consistently faster in task 1.

McGonigle and Chalmers,The Ontology of Order, 1996

Affirmativity and monotonicity?

- Monotonicity is a semantic property of quantifiers;
- Degree of affirmativity is a linguistic concept, e.g.,:
  - tag test;
  - licensing NPIs.

Example

1. Few children are dirty, are they?
2. Few children believe that any more.
3. *A few children are dirty, are they?
4. *A few children believe that any more.

A partial dissociation

Observation

Dissociation between downward monotonicity and negativity.

Example

1. At most half of the children believe that, don't they?
2. Not many children believe that, do they?

Observation

Downward monotone quantifiers fall into two classes: affirmatives and negatives.

The confound

Question

Are effects reported by JC'71 due to monotonicity or negativity?

Question

Does the data supports Bawise and Cooper's hypothesis?

Outline

Comparison Model

4 stage processing of the comparison model:

1. sentence encoding,
2. picture encoding,
3. comparing,
4. responding.

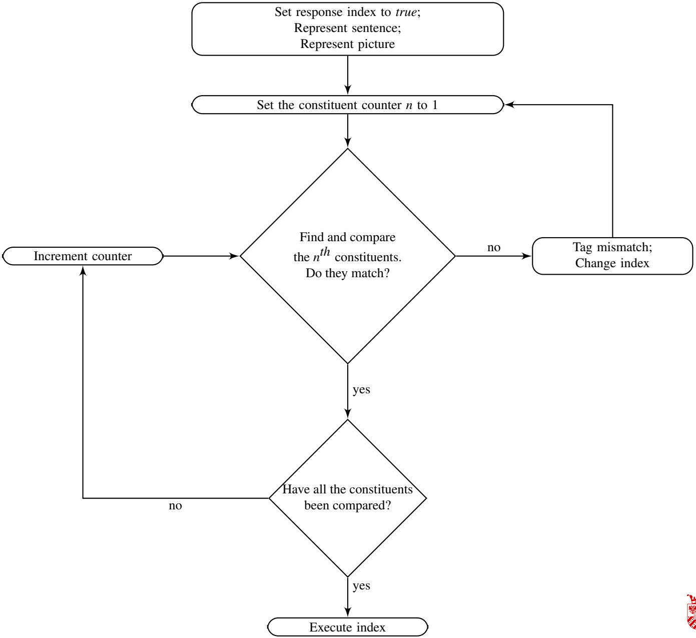
Component latencies are additive.

Clark & Chase, On the Process of Comparing Sentences Against Pictures, Cognitive Psychology, 1972

Example: The Dots are (not) White.

True affirmative			False affirmative		
Comparison:		Index	Comparison:		Index
Aff	(*white*, *dots*)		*Aff*	(*white*, *dots*)	
Aff	(*white*, *dots*)		*Aff*	(*black*, *dots*)	
+	+			-	*false*
		true	+	+	
					false
2 comparisons			3 comparisons		
True negative			**False negative**		
Comparison:		Index	Comparison:		Index
Neg	(*white*, *dots*)		*Neg*	(*white*, *dots*)	
Aff	(*white*, *dots*)		*Aff*	(*black*, *dots*)	
-	+	*false*		-	*false*
+	+		-	+	*true*
			+	+	
4 comparisons			5 comparisons		

Computational Model



Pros & Cons of the Comparison Model

Pros

- *Explains variations w.r.t. negativity:*
  - *negatives harder than affirmatives;*
  - *true affirmatives easier than false ones;*
  - *false negatives easier than true ones.*

Cons

- *It says very little about monotonicity.*
- *Arbitrary psychological/linguistic representation.*
- *Little insight into the actual computational process.*
- *At best, post-hoc theory for the specific task.*

Tanenhaus et al., Sentence-Picture Verification Models as Theories of Sentence Comprehension: A Critique of Carpenter and Just, Psychological Review, 1976

Question

How to cover generalized quantifiers?