Delimited continuations in natural language: Quantification and polarity sensitivity

Chung-chieh Shan
Harvard University
ccshan@post.harvard.edu

Continuations Workshop
17 January 2004

$$\begin{tikzcd}[column sep=huge] \text{Computational} \arrow[r, bend right] & \text{Linguistics} \arrow[l, bend right] \end{tikzcd}$$

What a linguist cares about

Entailment

every student enjoyed a conference	$\nvdash$	every student enjoyed POPL
no student enjoyed a conference	$\vdash$	no student enjoyed POPL
a student enjoyed a conference	$\nvdash$	a student enjoyed POPL
most students enjoyed a conference	$\nvdash$	most students enjoyed POPL

Ambiguity

Did some student enjoy every conference?

$$\exists x.\, \forall y.\, \mathsf{enjoyed}(x, y)$$
$$\forall y.\, \exists x.\, \mathsf{enjoyed}(x, y)$$

Did any student enjoy every conference?

Acceptability

*every student enjoyed any conference
no student enjoyed any conference
*a student enjoyed any conference
*most students enjoyed any conference

This talk deals with English, but the approach hopefully extends to other languages (which are different!).

Translation to a logical metalanguage

$$\begin{array}{ccc}
\text{Every student enjoyed a conference} & \not\models & \text{Every student enjoyed POPL} \\
\downarrow & & \downarrow \\
\forall x.\, \mathsf{student}(x) \Rightarrow \exists y.\, \mathsf{conf}(y) \wedge \mathsf{enjoyed}(x, y) & \not\models & \forall x.\, \mathsf{student}(x) \Rightarrow \mathsf{enjoyed}(x, \mathsf{popl}) \\
\downarrow & & \downarrow \\
\langle \text{some truth condition on models} \rangle & \not\models & \langle \text{some truth condition on models} \rangle
\end{array}$$

The guiding analogy

Programming languages	**Natural languages**
desired behavior	speaker judgments
observations at ground type	truth conditions, etc.
type system	syntax
denotational semantics	denotational semantics
computational side effects	“linguistic side effects”
control effects	quantification, polarity, etc.

Computational side effects in the logical metalanguage . . .
. . . handles *“linguistic side effects”*

State in the logical metalanguage . . .
. . . handles pronouns and binding

Control operators in the logical metalanguage . . .
. . . handles quantification and ***polarity sensitivity***

Outline

- ✓ Overview
- ▶ A simple grammatical formalism
- Quantification with shift and reset
- Quantifier scope ambiguity
- Polarity sensitivity

Computational side effects in the logical metalanguage . . .
. . . handles *“linguistic side effects”*

State in the logical metalanguage . . .
. . . handles pronouns and binding

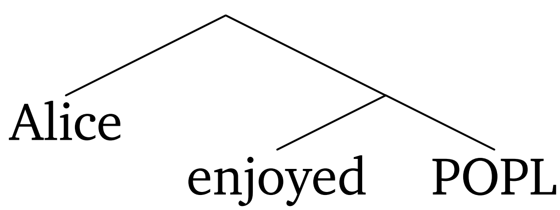
Control operators in the logical metalanguage . . .
. . . handles quantification and ***polarity sensitivity***

A simple grammatical formalism

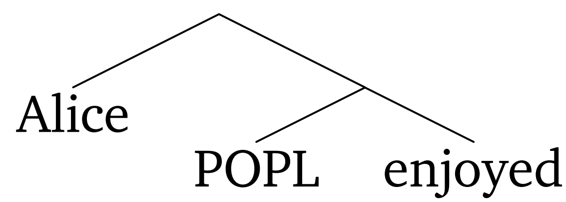
Alice enjoyed POPL. *Alice enjoyed. *Alice enjoyed Bob POPL.

$$
\begin{aligned}
[\![\text{Alice}]\!] &= \mathsf{alice} && : \mathsf{Thing} \\
[\![\text{POPL}]\!] &= \mathsf{popl} && : \mathsf{Thing} \\
[\![\text{enjoyed}]\!] &= \mathsf{enjoyed} && : \mathsf{Thing} \xrightarrow{\prime} (\mathsf{Thing} \xrightarrow{\grave{}} \mathsf{Bool})
\end{aligned}
$$

$$
\begin{aligned}
f \mathbin{/} x &= f(x) && : \beta && \text{where } f : \alpha \xrightarrow{\prime} \beta, \quad x : \alpha \\
x \mathbin{\backslash} f &= f(x) && : \beta && \text{where } f : \alpha \xrightarrow{\grave{}} \beta, \quad x : \alpha
\end{aligned}
$$



$\mathsf{alice} \mathbin{\backslash} (\mathsf{enjoyed} \mathbin{/} \mathsf{popl}) : \mathsf{Bool}$



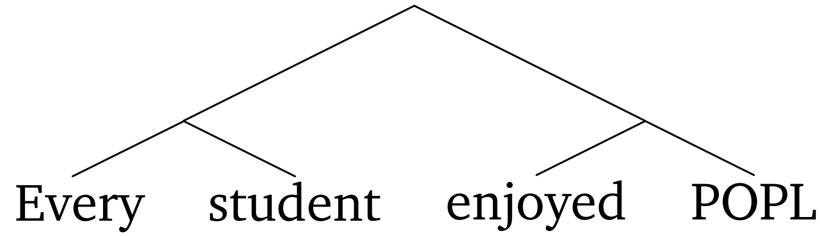
$\cancel{\mathsf{alice} \mathbin{\backslash} (\mathsf{popl} \mathbin{\backslash} \mathsf{enjoyed}) : \mathsf{Bool}}$

Right-associative by convention.

Notational variant of combinatory categorial grammar.

Quantification

Every student enjoyed POPL. $\forall x.\, \mathsf{student}(x) \Rightarrow \mathsf{enjoyed}(\mathsf{popl})(x)$

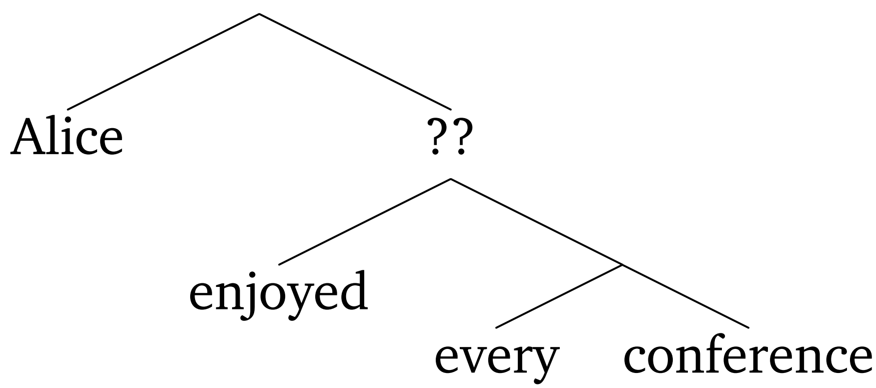


$[\![\text{Every student}]\!] = ??$

$$\mathsf{student} : \mathsf{Thing} \to \mathsf{Bool}$$

$$[\![\text{every}]\!] = \lambda r.\, \lambda s.\, \forall x.\, r(x) \Rightarrow s(x) : (\mathsf{Thing} \to \mathsf{Bool}) \acute{\to} (\mathsf{Thing} \grave{\to} \mathsf{Bool}) \acute{\to} \mathsf{Bool}$$

$$[\![\text{some}]\!] = \lambda r.\, \lambda s.\, \exists x.\, r(x) \wedge s(x) : (\mathsf{Thing} \to \mathsf{Bool}) \acute{\to} (\mathsf{Thing} \grave{\to} \mathsf{Bool}) \acute{\to} \mathsf{Bool}$$



Quantification with shift and reset

We want:

$$[\![\text{every conference}]\!] = \xi s.\, \forall x.\, \mathsf{conf}(x) \Rightarrow s(x) : \mathsf{Thing}_{\mathsf{Bool}}^{\mathsf{Bool}}.$$

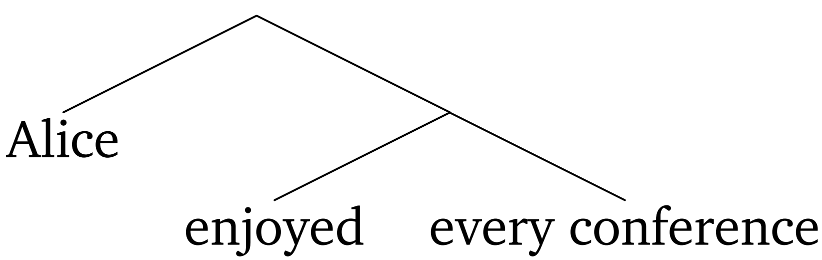
Here the type $\mathsf{Thing}_{\mathsf{Bool}}^{\mathsf{Bool}}$ has the CPS transform $(\mathsf{Thing} \to \mathsf{Bool}) \to \mathsf{Bool}$.

In general, α_{γ}^{δ} has the CPS transform $(\alpha \to \gamma) \to \delta$.

Reverse-engineer denotations for "every" and "some":

$$[\![\text{every}]\!] = \lambda r.\, \xi s.\, \forall x.\, r(x) \Rightarrow s(x) : (\mathsf{Thing} \to \mathsf{Bool}) \mathrel{\acute{\to}} \mathsf{Thing}_{\mathsf{Bool}}^{\mathsf{Bool}}$$
$$[\![\text{some}]\!] = \lambda r.\, \xi s.\, \exists x.\, r(x) \wedge s(x) : (\mathsf{Thing} \to \mathsf{Bool}) \mathrel{\acute{\to}} \mathsf{Thing}_{\mathsf{Bool}}^{\mathsf{Bool}}$$

Can now handle quantificational noun phrases in any position:



$$\big[[\![\text{Alice}]\!] \backslash [\![\text{enjoyed}]\!] / [\![\text{every conference}]\!]\big]$$
$$= [\mathsf{alice} \backslash \mathsf{enjoyed} / \underline{\xi s.\, \forall x.\, \mathsf{conf}(x) \Rightarrow s(x)}]$$
$$\rhd [\forall x.\, \mathsf{conf}(x) \Rightarrow \underline{(\lambda v.\, [\mathsf{alice} \backslash \mathsf{enjoyed} / v])(x)}]$$
$$\rhd \cdots$$
$$\rhd \forall x.\, \mathsf{conf}(x) \Rightarrow \mathsf{alice} \backslash \mathsf{enjoyed} / x : \mathsf{Bool}$$

Notion of evaluation! Beginnings of psycholinguistics.

Quantifier scope ambiguity

Nondeterminism in natural language:

Some student enjoyed every conference.

$$\begin{array}{ll} \exists x.\, \forall y.\, \mathsf{enjoyed}(y)(x) & \leftarrow \text{linear scope} \\ \forall y.\, \exists x.\, \mathsf{enjoyed}(y)(x) & \leftarrow \text{inverse scope} \end{array}$$

How to generate ambiguity?

$$\begin{array}{ll} & \text{Deterministic composition} \\ + & \text{Deterministic word meanings} \\ \hline & \text{Deterministic sentence meanings} \end{array}$$

Two approaches:

- ***Nondeterministic evaluation order***
  Quantifiers evaluated earlier scope wider.

- ▶ ***Hierarchy of control operators***
  Quantifiers at outer levels scope wider.
  (cf. TDPE paper at this POPL by Balat, Di Cosmo, and Fiore)
  People tend to process words in the order they are spoken.

Pronouns, questions, and polarity favor the second approach, but it needs staging—

Quantifier scope ambiguity with hierarchy & staging

$$
\begin{aligned}
&\big[[\![\text{some student}]\!] \backslash [\![\text{enjoyed}]\!] / [\![\text{every conference}]\!]\big] \\
&= \Big[\big(\underline{\xi^2 s.\ \exists x.\ \mathsf{student}(x) \wedge s(x)}\big) \backslash \mathsf{enjoyed} / \big(\xi^1 t.\ \forall y.\ \mathsf{conf}(y) \Rightarrow t(y)\big)\Big]^0 \\
&\rhd \Big[\exists x.\ \mathsf{student}(x) \wedge \underline{\big(\lambda v.\ [v \backslash \mathsf{enjoyed} / \xi^1 t.\ \forall y.\ \mathsf{conf}(y) \Rightarrow t(y)]^2\big)(x)}\Big]^0 \\
&\rhd \cdots \\
&\rhd \Big[\exists x.\ \mathsf{student}(x) \wedge [x \backslash \mathsf{enjoyed} / \underline{\xi^1 t.\ \forall y.\ \mathsf{conf}(y) \Rightarrow t(y)}]^2\Big]^0 \\
&\rhd \Big[\forall y.\ \mathsf{conf}(y) \Rightarrow \underline{\big(\lambda v.\ [\exists x.\ \mathsf{student}(x) \wedge [x \backslash \mathsf{enjoyed} / v]^2]^1\big)(y)}\Big]^0 \\
&\rhd \Big[\forall y.\ \mathsf{conf}(y) \Rightarrow [\exists x.\ \mathsf{student}(x) \wedge [x \backslash \mathsf{enjoyed} / y]^2]^1\Big]^0
\end{aligned}
$$

What is $\exists$ above, really?

- Is it ***higher-order abstract syntax*** : $(\mathsf{Thing} \to \mathsf{Bool}) \to \mathsf{Bool}$?
  No, because then the body under $\exists$ must be pure.
- Is it ***gensym and first-order abstract syntax*** : $\mathsf{Bool} \to \mathsf{Bool}$?
  Perhaps, but need to rule out unbound x in $\exists x.\ \xi f.\ \mathsf{student}(x) \rhd \mathsf{student}(x)$.
  (cf. "Some student enjoyed every conference s/he organized.")
- Ideally, it is higher-order abstract syntax ***staged*** in a language with control.

Two kinds of control hierarchies

Danvy and Filinski:

- Post-CPS types look like $\alpha^{\left(\gamma_0{}^{\delta_0}_{\delta_1}\right)}_{\left(\gamma_1{}^{\delta_2}_{\delta_3}\right)}$
- Needs gensym for now

Barker and Shan:

- Post-CPS types look like $\left(\alpha^{\gamma_0}_{\gamma_1}\right)^{\delta_0}_{\delta_1}$
- No direct-style terms yet

Both improve Hobbs and Shieber's and Lewin's quantifier scoping algorithms:

- ✓ Directly compositional, not a post-processing step after parsing
- ✓ Semantically motivated by delimited continuations
- Interacts properly with other linguistic side effects
  - ✓ (other) quantification
  - pronouns
  - questions

Outline

- ✓ Overview
- ✓ A simple grammatical formalism
- ✓ Quantification with shift and reset
- ✓ Quantifier scope ambiguity
- ▶ Polarity sensitivity

Computational side effects in the logical metalanguage …
… handles *"linguistic side effects"*

State in the logical metalanguage …
… handles pronouns and binding

Control operators in the logical metalanguage …
… handles quantification and ***polarity sensitivity***

Polarity sensitivity

The quantifiers "a", "some", and "any"† all look existential:

Did a student call?
Did some student call?
Did any student call?

$$\exists x.\, \mathsf{student}(x) \wedge \mathsf{called}(x)$$

But do not behave the same:

No student enjoyed some conference.	(unambiguous $\exists\neg$)
No student enjoyed a conference.	(ambiguous $\neg\exists$, $\exists\neg$)
No student enjoyed any conference.	(unambiguous $\neg\exists$)
Some student enjoyed no conference.	(unambiguous $\exists\neg$)
A student enjoyed no conference.	(ambiguous $\neg\exists$, $\exists\neg$)
*Any student enjoyed no conference.	(unacceptable)

"Any" is a ***negative polarity item***:
Very roughly, it requires negative contexts, such as in the scope of "no".

"Some" is a ***positive polarity item***:
Very roughly, it is allergic to negative contexts.

Meaning affects ambiguity and acceptability! But linear order matters too.

Chaining answer types

What an under-appreciated feature of shift and reset!

$$\frac{\Gamma \vdash F : (\alpha \xrightarrow{\prime} \beta^{\gamma_2}_{\gamma_3})^{\gamma_0}_{\gamma_1} \qquad \Gamma \vdash E : \alpha^{\gamma_1}_{\gamma_2}}{\Gamma \vdash F \,/\, E : \beta^{\gamma_0}_{\gamma_3}} \xrightarrow{\prime}\mathrm{E}$$

$$\frac{\Gamma \vdash E : \alpha^{\gamma_0}_{\gamma_1} \qquad \Gamma \vdash F : (\alpha \xrightarrow{\grave{}} \beta^{\gamma_2}_{\gamma_3})^{\gamma_1}_{\gamma_2}}{\Gamma \vdash E \,\backslash\, F : \beta^{\gamma_0}_{\gamma_3}} \xrightarrow{\grave{}}\mathrm{E}$$

$$\begin{tikzcd}
\cdots \arrow[r, bend left] & \gamma_0 \arrow[r, bend left, "\gamma_0 \atop \gamma_1"] & \gamma_1 \arrow[r, bend left, "\gamma_1 \atop \gamma_2"] & \gamma_2 \arrow[r, bend left, "\gamma_2 \atop \gamma_3"] & \gamma_3 \arrow[r, bend left] & \cdots
\end{tikzcd}$$

Polarity sensitivity with answer-type subtyping

A standard approach to modeling polarity sensitivity:
split the answer type Bool into a family of subtypes.

$$\mathsf{Bool} \leq \mathsf{BoolPos} \qquad \mathsf{Bool} \leq \mathsf{BoolNeg}$$

(in addition to the usual rules for subtyping)

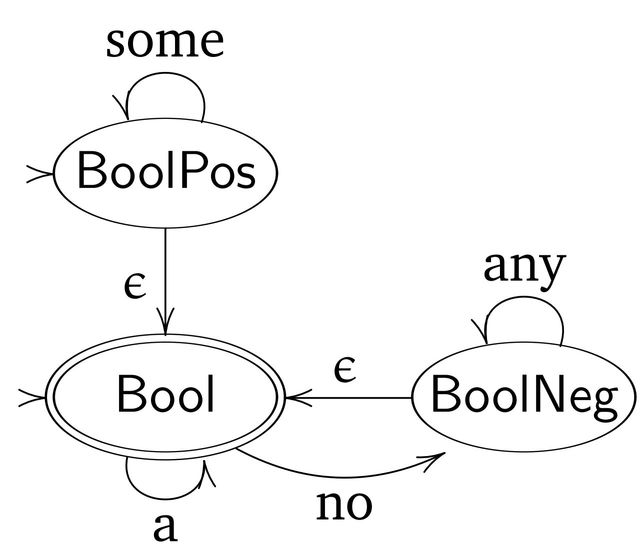
The return type of verbs like "enjoyed" remains Bool.

Also, restrict Reset to produce the answer type Bool or BoolPos, not BoolNeg.

$$\frac{\Gamma \vdash E : \alpha_{\alpha}^{\beta}}{\Gamma \vdash [E] : \beta}\,\text{Reset} \qquad \text{where } \beta \leq \mathsf{BoolPos}$$

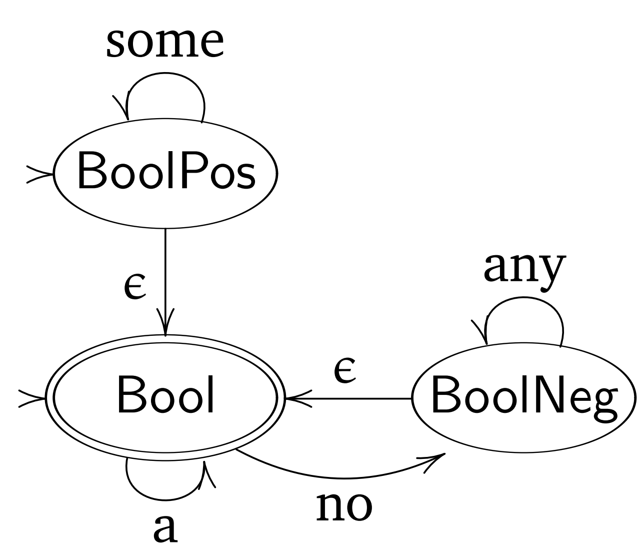
Finally, refine the answer types for quantifiers.

$$\llbracket \text{no} \rrbracket : (\mathsf{Thing} \to \mathsf{Bool}) \acute{\to} \mathsf{Thing}^{\mathsf{Bool}}_{\mathsf{BoolNeg}},$$
$$\llbracket \text{some} \rrbracket : (\mathsf{Thing} \to \mathsf{Bool}) \acute{\to} \mathsf{Thing}^{\mathsf{BoolPos}}_{\mathsf{BoolPos}},$$
$$\llbracket \text{a} \rrbracket : (\mathsf{Thing} \to \mathsf{Bool}) \acute{\to} \mathsf{Thing}^{\mathsf{Bool}}_{\mathsf{Bool}},$$
$$\llbracket \text{any} \rrbracket : (\mathsf{Thing} \to \mathsf{Bool}) \acute{\to} \mathsf{Thing}^{\mathsf{BoolNeg}}_{\mathsf{BoolNeg}}.$$



Polarity sensitivity: revisiting empirical data

Quantifiers with levels	Type	Reading
$\text{no}^1 \ldots \text{some}^2$	$: \text{Bool}^{\text{BoolPos}^{\text{Bool}}_{\text{BoolNeg}}}_{\text{BoolPos}\cdots}$	$\not\to$
$\text{no}^1 \ldots \text{some}^1$	$:/$	
$\text{no}^2 \ldots \text{some}^1$	$: \text{Bool}^{\text{Bool}^{\text{BoolPos}}_{\text{BoolPos}}}_{\text{BoolNeg}\cdots}$	$\to (\exists\neg)$
$\text{no}^1 \ldots \text{a}^2$	$: \text{Bool}^{\text{Bool}^{\text{Bool}}_{\text{BoolNeg}}}_{\text{Bool}\cdots}$	$\to (\neg\exists)$
$\text{no}^1 \ldots \text{a}^1$	$: \text{Bool}^{\text{Bool}}_{\text{Bool}}$	$\to (\neg\exists)$
$\text{no}^2 \ldots \text{a}^1$	$: \text{Bool}^{\text{Bool}^{\text{Bool}}_{\text{Bool}}}_{\text{BoolNeg}\cdots}$	$\to (\exists\neg)$
$\text{no}^1 \ldots \text{any}^2$	$: \text{Bool}^{\text{BoolNeg}^{\text{Bool}}_{\text{BoolNeg}}}_{\text{BoolNeg}\cdots}$	$\not\to$
$\text{no}^1 \ldots \text{any}^1$	$: \text{Bool}^{\text{Bool}}_{\text{BoolNeg}}$	$\to (\neg\exists)$
$\text{no}^2 \ldots \text{any}^1$	$: \text{Bool}^{\text{Bool}^{\text{BoolNeg}}_{\text{BoolNeg}}}_{\text{BoolNeg}\cdots}$	$\not\to$

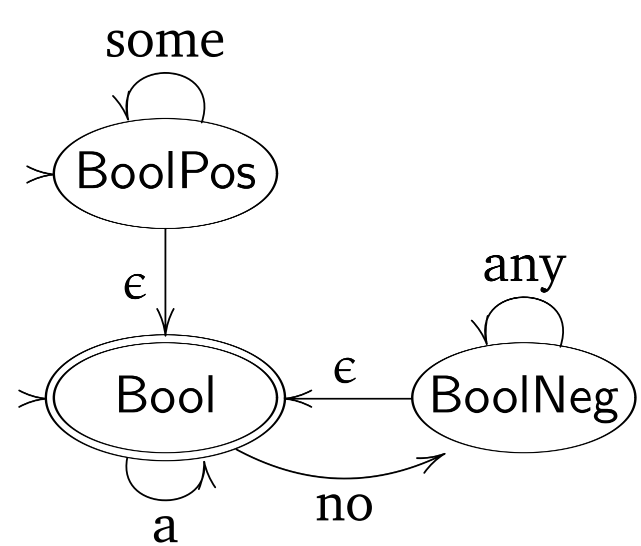


"no . . . any" must be on the same level
"no" must scope over "some"

Polarity sensitivity: revisiting empirical data

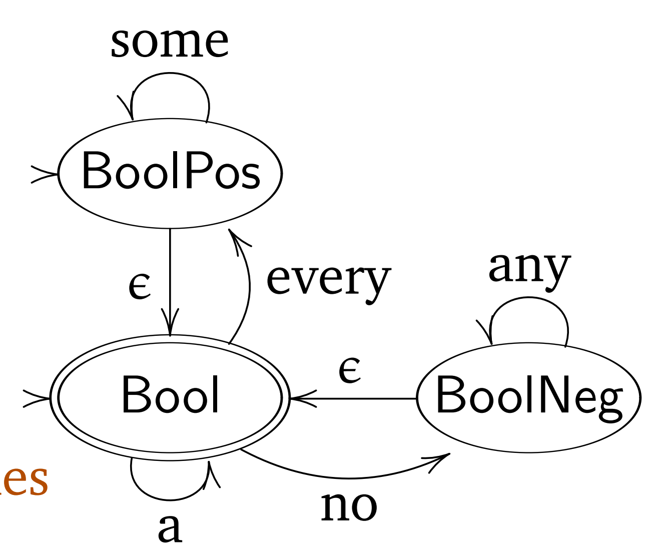
Quantifiers with levels	Type	Reading
$\text{some}^1 \ldots \text{no}^2$	$: \text{Bool}^{\text{Bool}^{\text{BoolPos}}_{\text{BoolPos}}}_{\text{BoolNeg}\cdots}$	$\longrightarrow (\exists\neg)$
$\text{some}^1 \ldots \text{no}^1$	$: \text{Bool}^{\text{BoolPos}}_{\text{BoolNeg}}$	$\longrightarrow (\exists\neg)$
$\text{some}^2 \ldots \text{no}^1$	$: \text{Bool}^{\text{BoolPos}^{\text{Bool}}_{\text{BoolNeg}}}_{\text{BoolPos}\cdots}$	$\not\longrightarrow$
$\text{a}^1 \ldots \text{no}^2$	$: \text{Bool}^{\text{Bool}^{\text{Bool}}_{\text{Bool}}}_{\text{BoolNeg}\cdots}$	$\longrightarrow (\exists\neg)$
$\text{a}^1 \ldots \text{no}^1$	$: \text{Bool}^{\text{Bool}}_{\text{BoolNeg}}$	$\longrightarrow (\exists\neg)$
$\text{a}^2 \ldots \text{no}^1$	$: \text{Bool}^{\text{Bool}^{\text{Bool}}_{\text{BoolNeg}}}_{\text{Bool}\cdots}$	$\longrightarrow (\neg\exists)$
$\text{any}^1 \ldots \text{no}^2$	$: \text{Bool}^{\text{Bool}^{\text{BoolNeg}}_{\text{BoolNeg}}}_{\text{BoolNeg}\cdots}$	$\not\longrightarrow$
$\text{any}^1 \ldots \text{no}^1$	$: \text{Bool}^{\text{BoolNeg}}_{\text{Bool}}$	$\not\longrightarrow$
$\text{any}^2 \ldots \text{no}^1$	$: \text{Bool}^{\text{BoolNeg}^{\text{Bool}}_{\text{BoolNeg}}}_{\text{BoolNeg}\cdots}$	$\not\longrightarrow$

“no . . . any” must be on the same level, in that order
“no” must scope over “some”



Polarity sensitivity: revisiting empirical data

Quantifiers with levels	Type	Reading
Every^1 student enjoyed some^1 conference	$: \mathsf{Bool}^{\mathsf{Bool}}_{\mathsf{BoolPos}} \longrightarrow$	$(\forall\exists)$
A^1 student enjoyed every^1 conference	$: \mathsf{Bool}^{\mathsf{Bool}}_{\mathsf{BoolPos}} \longrightarrow$	$(\exists\forall)$
$[\![\text{every}]\!] : (\mathsf{Thing} \to \mathsf{Bool}) \acute{\to} \mathsf{Thing}^{\mathsf{Bool}}_{\mathsf{BoolPos}}$		
No^1 professor gave every^1 student some^1 book	$: \mathsf{Bool}^{\mathsf{Bool}}_{\mathsf{BoolPos}} \longrightarrow$	$(\neg\forall\exists)$



"no . . . any" must be on the same level, in that order
"no" must scope over "some", except if "every" intervenes

Comparing approaches to scope ambiguity

Two approaches:

- Nondeterministic evaluation order: Quantifiers evaluated earlier scope wider.
- ▶ Hierarchy of control operators: Quantifiers at outer levels scope wider.
  More complex perhaps, but captures more empirical data.

Polarity items (and pronouns, and questions) favor the second approach:

No student enjoyed any conference. (unambiguous $\neg\exists$)
*Any student enjoyed no conference. (unacceptable)

The first semantic analysis of polarity items to capture the long-observed sensitivity to linear order.

Implemented in a substructural logic of delimited continuations,
using Richard Moot's theorem prover Grail for categorial grammar.
(Joint work with Chris Barker.)

Summary

Control operators in the logical metalanguage handles quantification and polarity sensitivity.

Ingredients of this analysis include

- a control hierarchy,
- staged generation of logical formulas,
- left-to-right evaluation order, and
- changing and chaining answer types.

The first semantic analysis of polarity items to capture linear order.

Beyond the λ-calculus: operational semantics for natural language?