Lexical Uncertainty

Leon Bergen

Hurford’s constraint

Hurford’s constraint (Hurford, 1972): When a disjunction is literally equivalent to one of its disjuncts, it is infelicitous to use the disjunction.

Explains why “Mary went to Paris or France” is infelicitous.

- This is equivalent to “Mary went to France.”
- In general, if B implies A, then “A or B” is equivalent to “A.” Therefore, “A or B” will be infelicitous.

Hurford’s constraint

We will consider embedded implicatures which violate Hurford’s constraint (Chierchia, Fox, & Spector, 2009):

a) Some of the students passed the test. -> The speaker knows that not all of the students passed.

b) Some or all of the students passed the test. -> The speaker is uncertain about whether all passed.

c) Some or most of the students passed the test. -> The speaker knows that not all passed, but is uncertain about whether most did.

Hurford’s constraint

The contrast between “Some or all” and “Some or most” (“SvM”) is a problem for most (all?) pragmatic theories:

- Both utterances are literally equivalent, and both have the same complexity/cost.
- Differences in literal content or cost are the only properties that differentiate utterances in the RSA model (or in any extant game-theoretic model).

From the perspective of these models, utterances which have the same literal content and cost are identical.

The existence of these implicatures has been used to argue against purely Gricean approaches to pragmatics (Chierchia, Fox, & Spector, 2009; Meyer, 2013).

Rational speech acts model (RSA)

Notation:

- u = utterance
- w = possible world
- S(u|w) = probability that the speaker will choose utterance u given the goal of communicating world w.
- L(w|u) = probability that the listener assigns to world w after hearing utterance u.
- $\mathcal{L}$ = lexicon, s.t. $\mathcal{L}(u,w) = \begin{cases} 0 & \text{if } w \notin [\![u]\!] \\ 1 & \text{if } w \in [\![u]\!] \end{cases}$
- P(w) = prior probability of w
- c(u) = cost of utterance u

RSA: Literal listener

Literal listener: $$L_0(w|u,\mathcal{L}) = \frac{\mathcal{L}(u,w)P(w)}{\sum_{w'}\mathcal{L}(u,w')P(w')}$$
$$\propto \mathcal{L}(u,w)P(w)$$

This listener interprets an utterance by filtering out any worlds that are literally incompatible with this utterance.

RSA: Strategic speaker

Speaker model: $S_n(u|w) \propto e^{\lambda U_n(u|w)}$

$\lambda > 0$ is the inverse-temperature, which controls the degree of rationality of the speaker.

The utility function is defined by:
$$U_n(u|w) = -\log(\frac{1}{L_{n-1}(w|u)}) - c(u)$$
$$= \log(L_{n-1}(w|u)) - c(u)$$

The speaker is more likely to choose utterances that will be interpreted correctly by the listener L_{n-1}.

RSA: Strategic listener

Strategic listener: $L_n(w|u) \propto P(w)S_n(u|w)$

This listener interprets an utterance by integrating their prior knowledge with the likelihood that the speaker S_n would choose the utterance given different states of the world.

Rational speech act model (RSA)

$$
\begin{array}{ll}
\dots & \\
\downarrow & \\
S_1 & \\
\downarrow & \\
L_1 & L_n(w|u) \propto P(w)S_n(u|w) \\
\downarrow & \\
S_1 & U_n(u|w) = -\log(\frac{1}{L_{n-1}(w|u)}) - c(u) \\
\downarrow & \\
L_0 & L_0(w|u,\mathcal{L}) \propto \mathcal{L}(u,w)P(w)
\end{array}
$$

Rational speech act model (RSA)

...

$\downarrow$

S_1

$\downarrow$

L_1 $\quad L_n(w|u) \propto P(w) S_n(u|w)$

$\downarrow$

S_1 $\quad U_n(u|w) = -\log(\frac{1}{L_{n-1}(w|u)}) - c(u)$

Fixed Lexicon

$\downarrow$

L_0 $\quad L_0(w|u, \mathcal{L}) \propto \mathcal{L}(u, w) P(w)$

Lexical Uncertainty Model

...

$$S_1$$

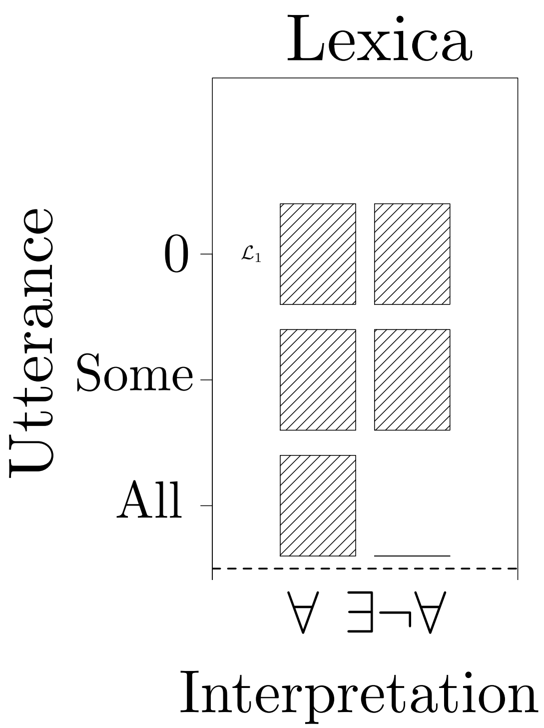
Infers the lexicon

$$L_1 \quad L_1(w|u,\mathcal{L}) \propto \sum_{\mathcal{L}'} P(w)P(\mathcal{L}')S_1(u|w,\mathcal{L}')$$

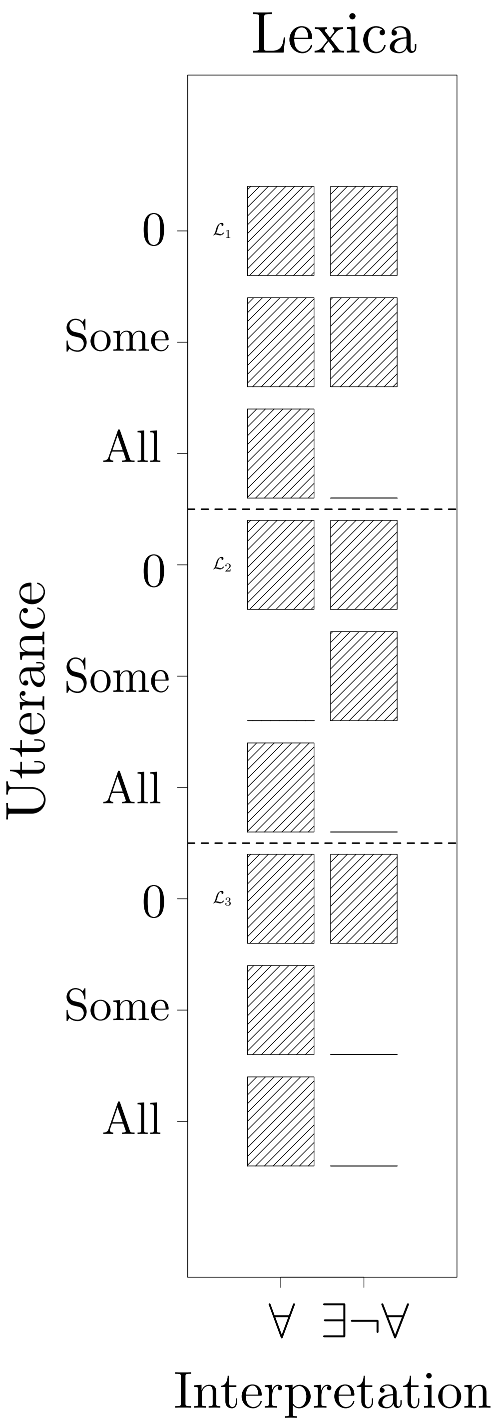
$$S_1 \quad U_1(u|w,\mathcal{L}) = -\log\frac{1}{L_0(w|u,\mathcal{L})} - c(u)$$

$$L_0 \quad L_0(w|u,\mathcal{L}) \propto \mathcal{L}(u,w)P(w)$$

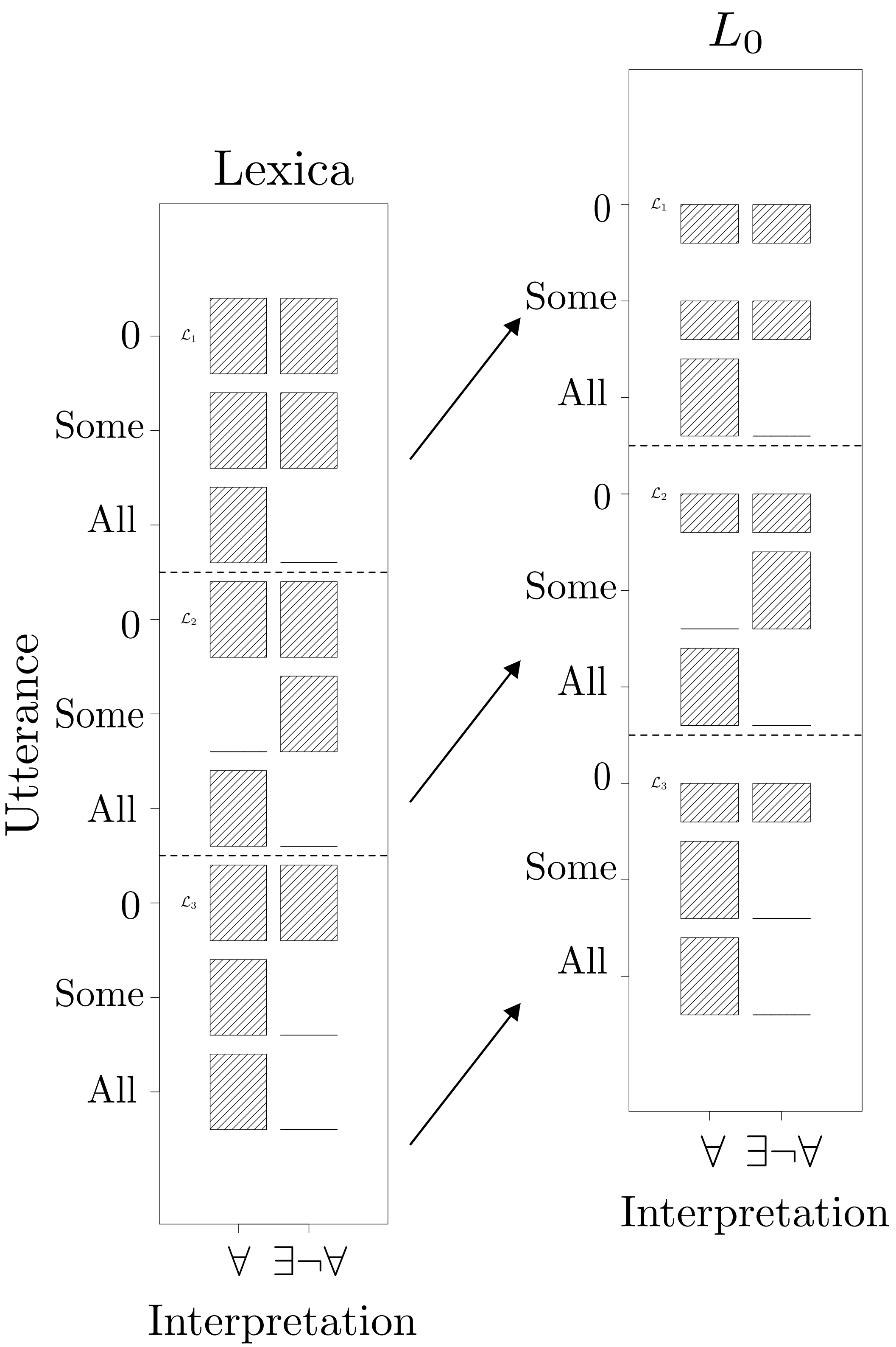
Derivation of scalar implicature



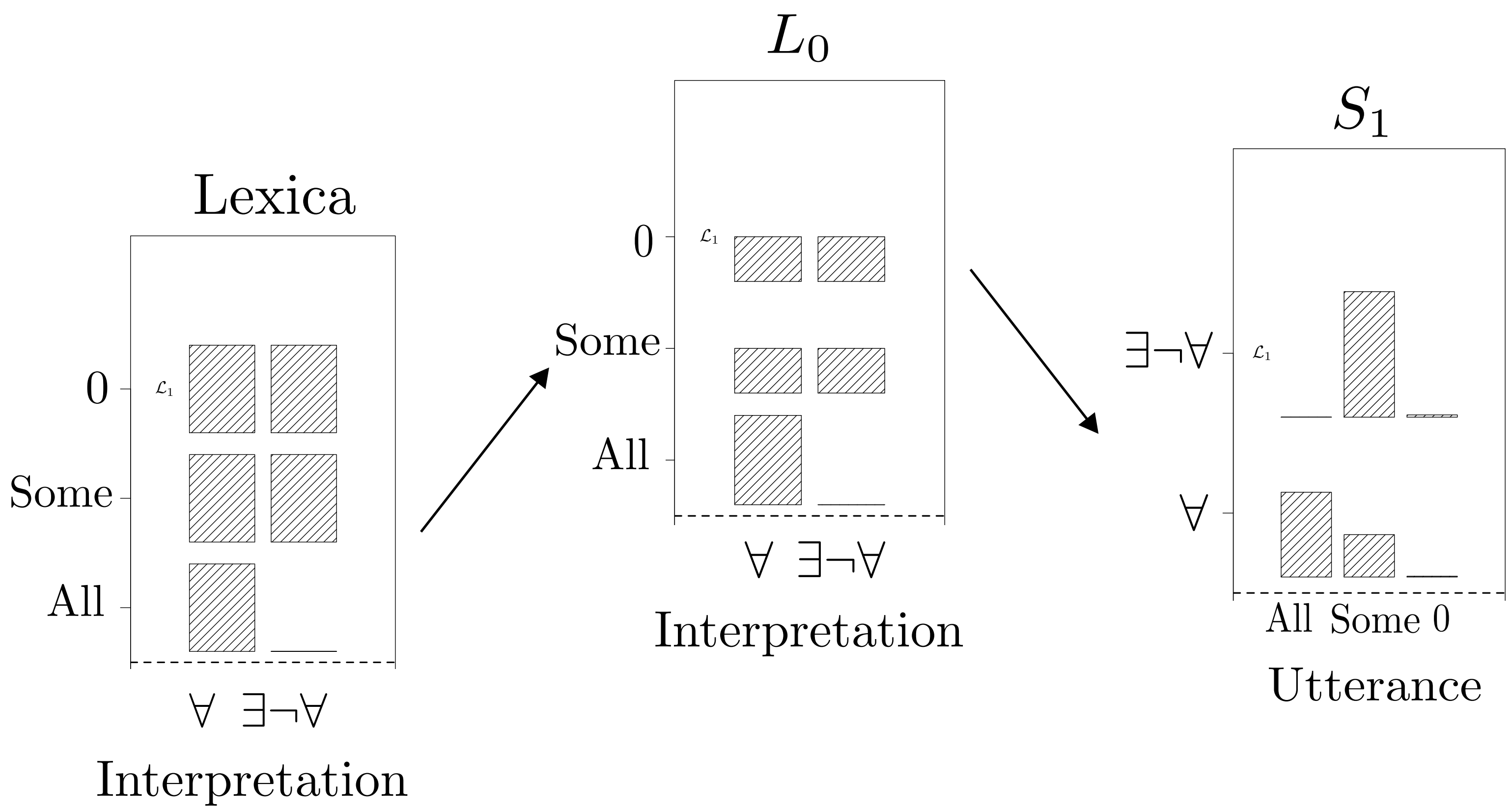
Derivation of scalar implicature



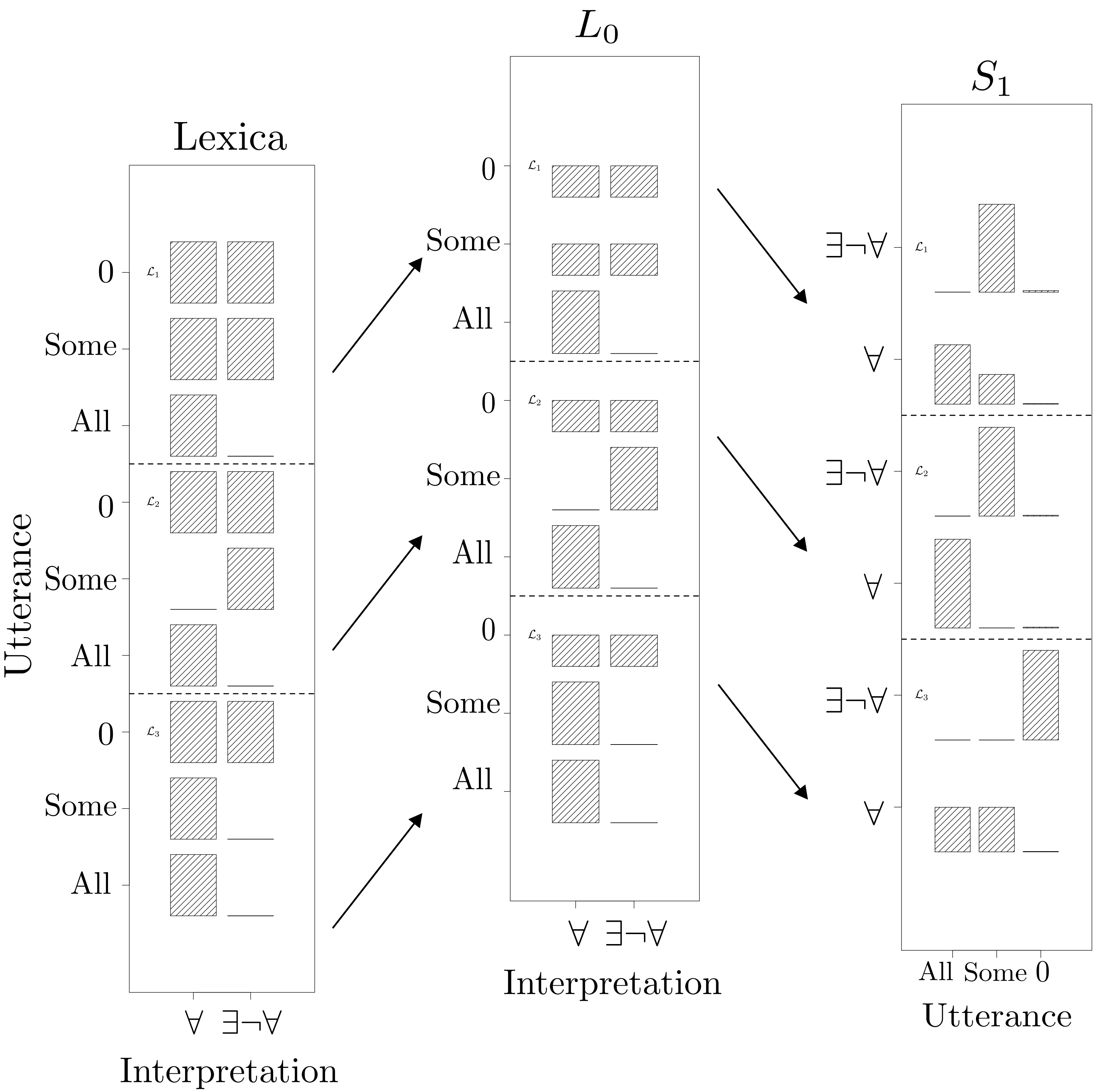
Derivation of scalar implicature



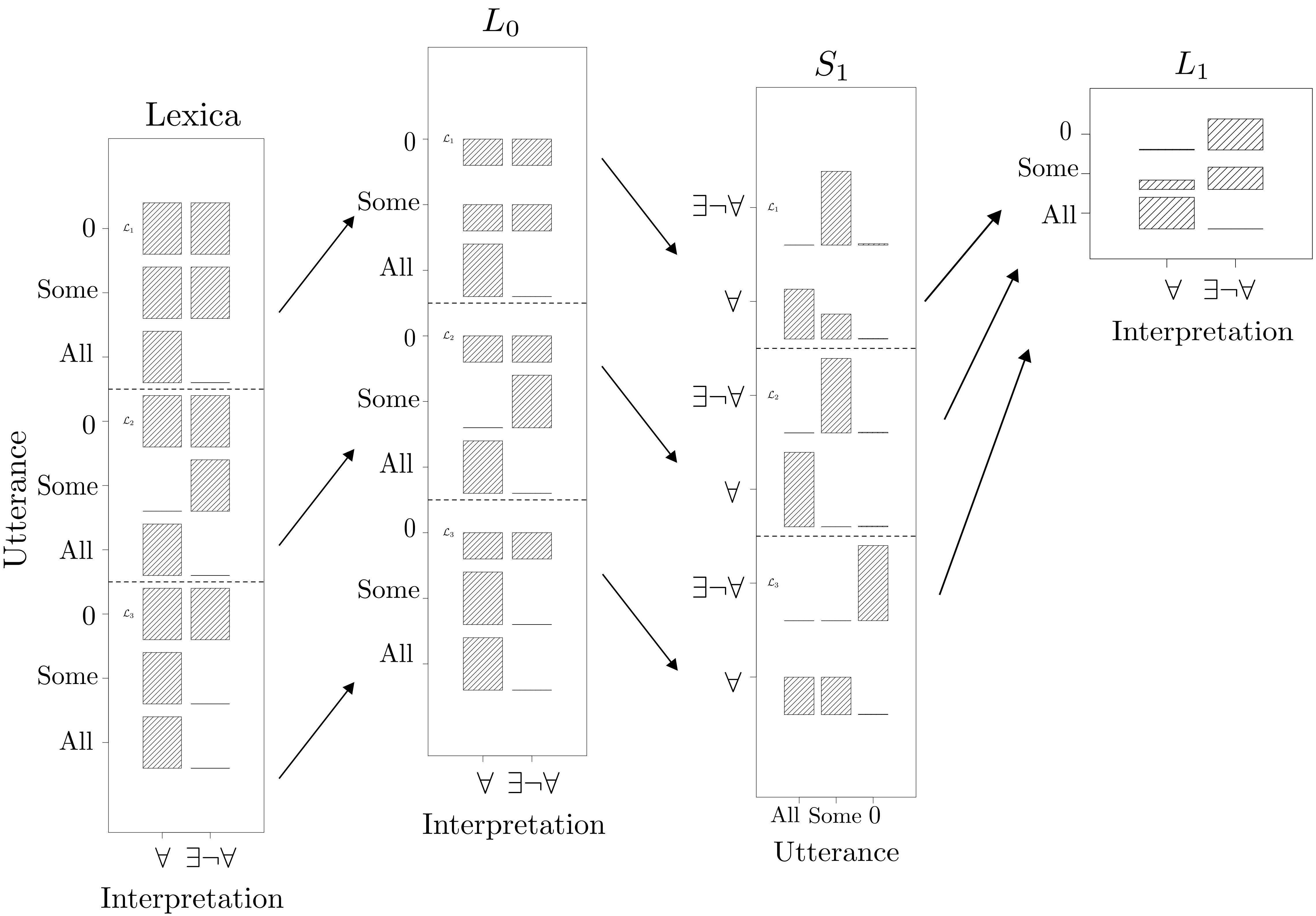
Derivation of scalar implicature



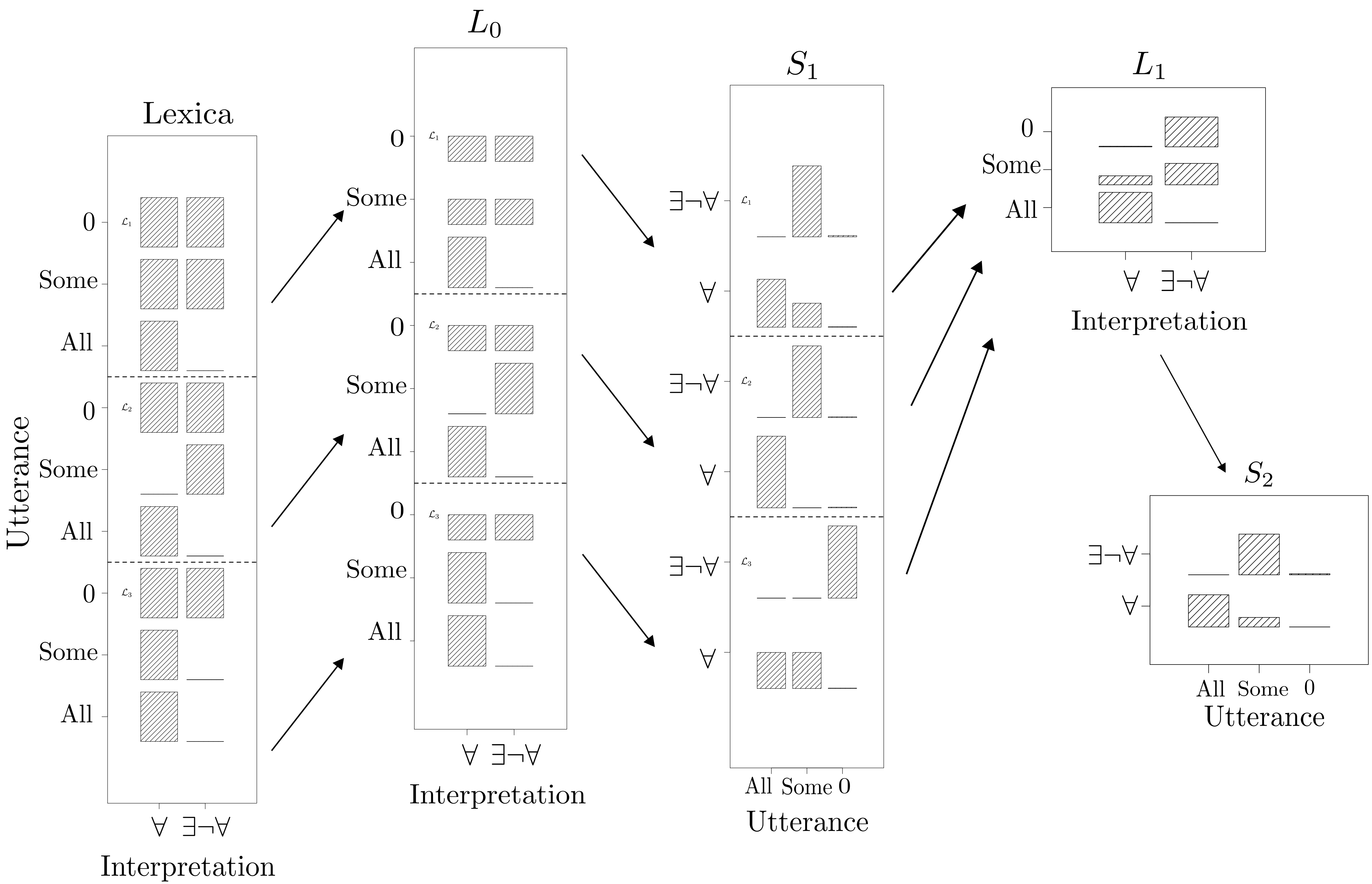
Derivation of scalar implicature



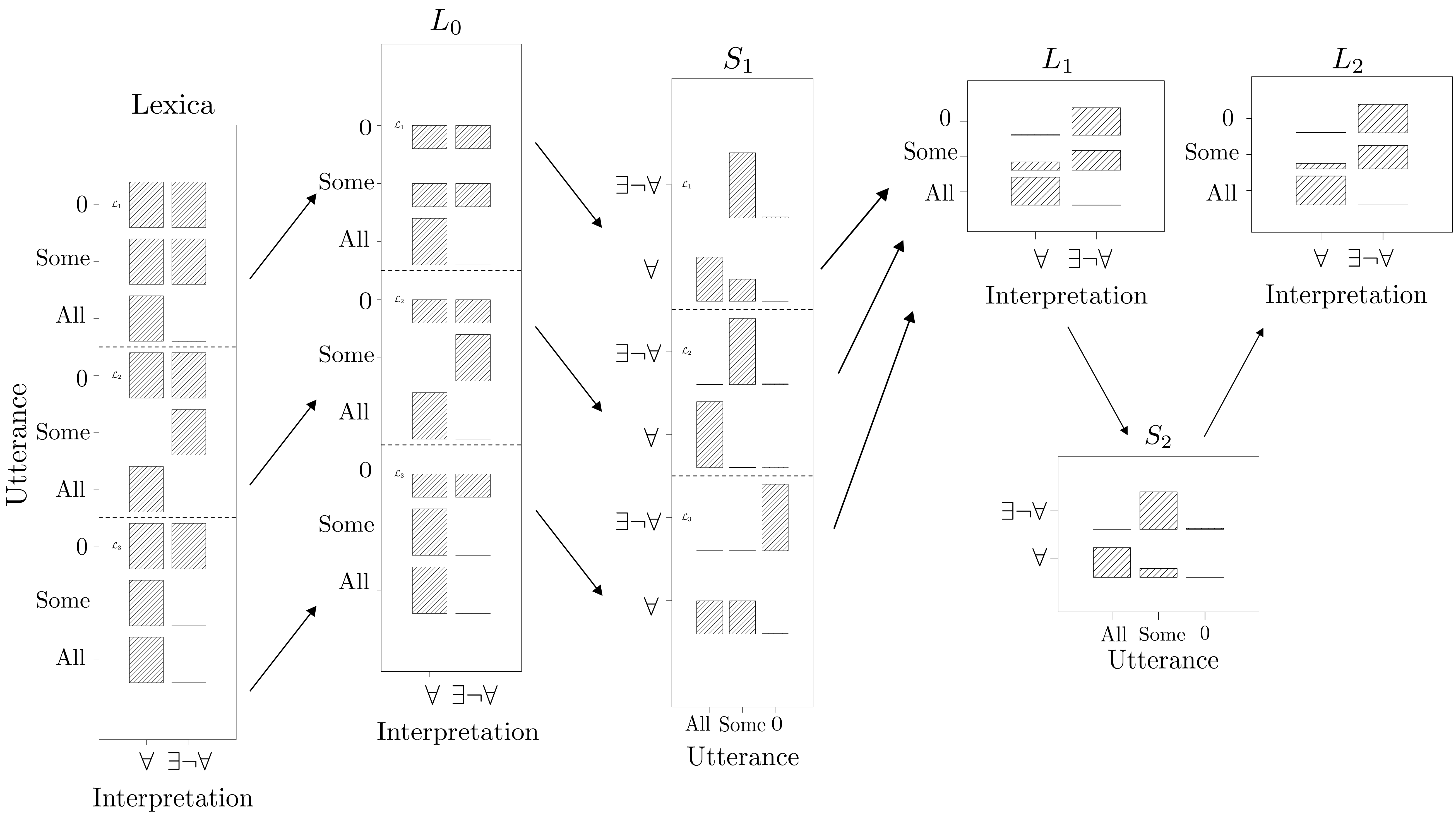
Derivation of scalar implicature



Derivation of scalar implicature



Derivation of scalar implicature



Non-convex disjunctive implicatures

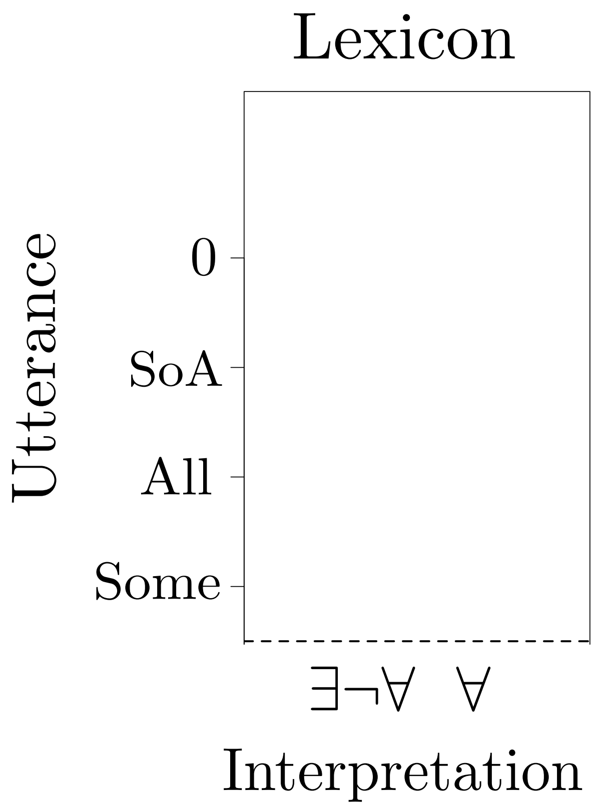
Context: A and B are visiting a resort but are frustrated with the temperature of the springs that they want to bathe in.

A: The springs in this resort are always warm or scalding.

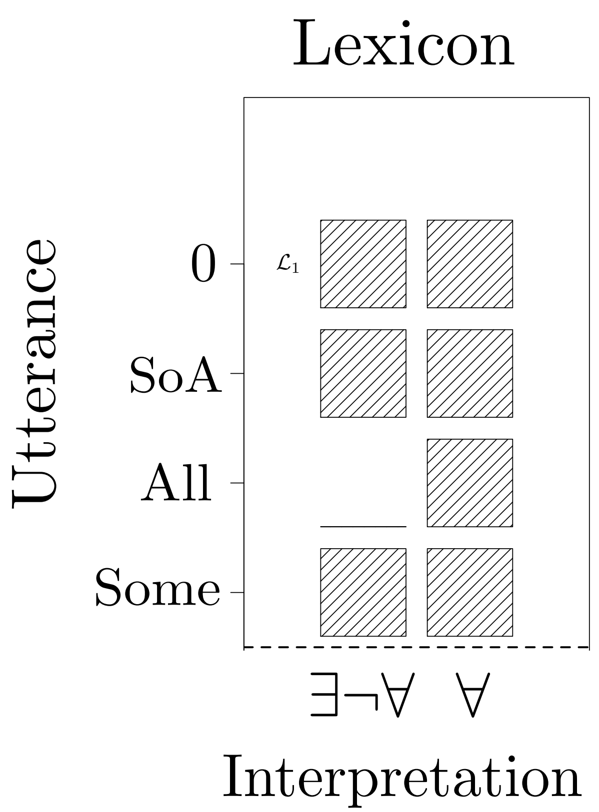
Context: A is discussing with B the performance of her son, who is extremely smart but blows off some classes, depending on how he likes the teacher.

A: My son's performance in next semester's math class will be adequate or stellar.

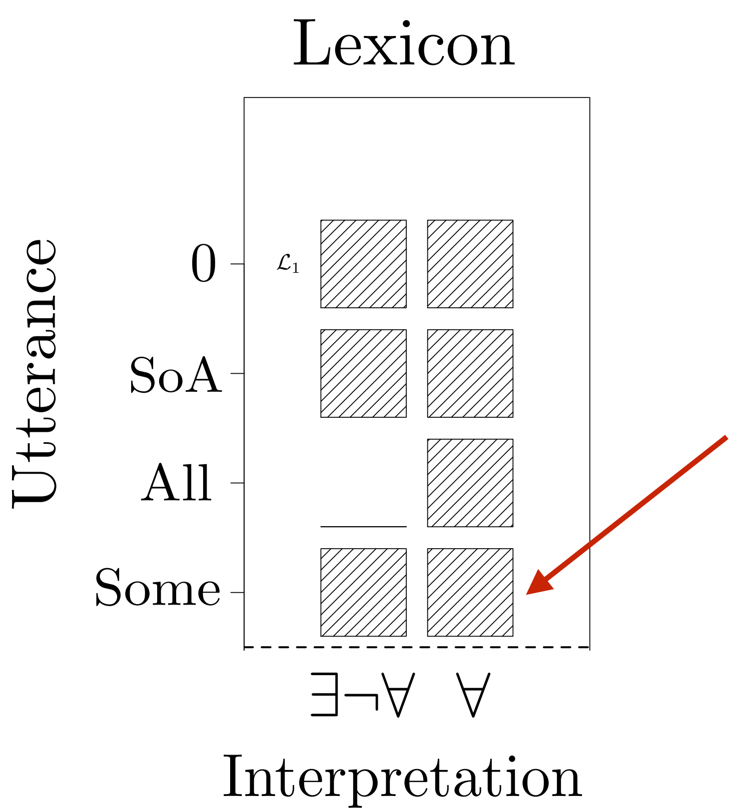
Derivation of embedded implicatures



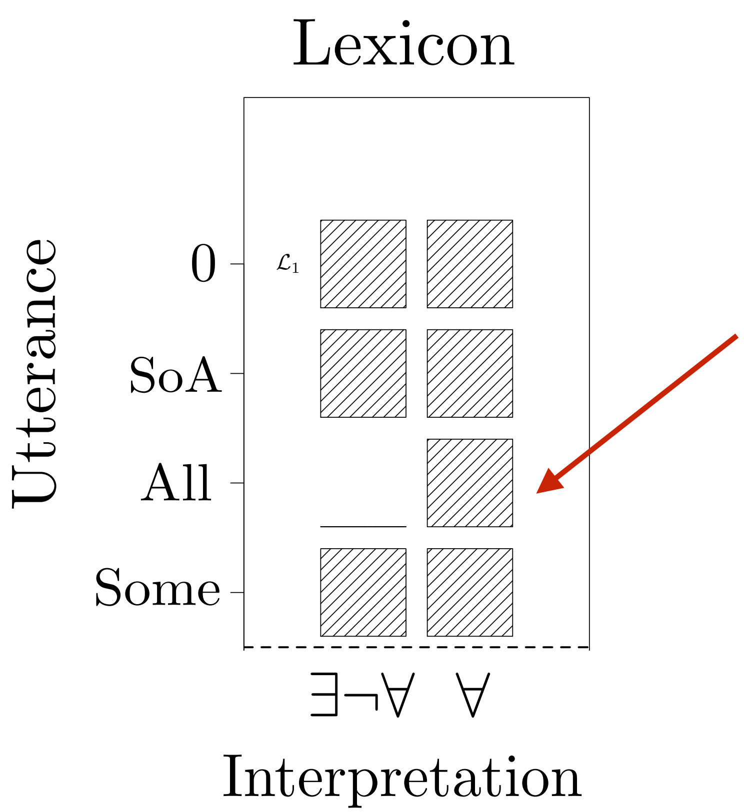
Derivation of embedded implicatures



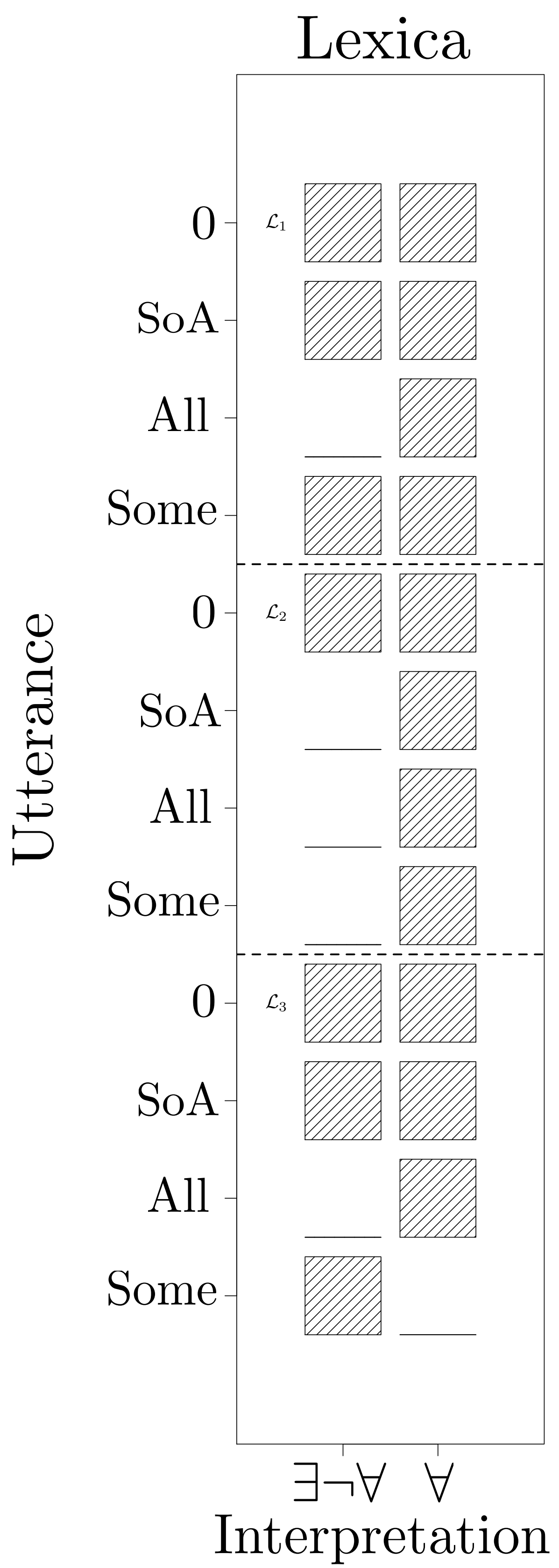
Derivation of embedded implicatures



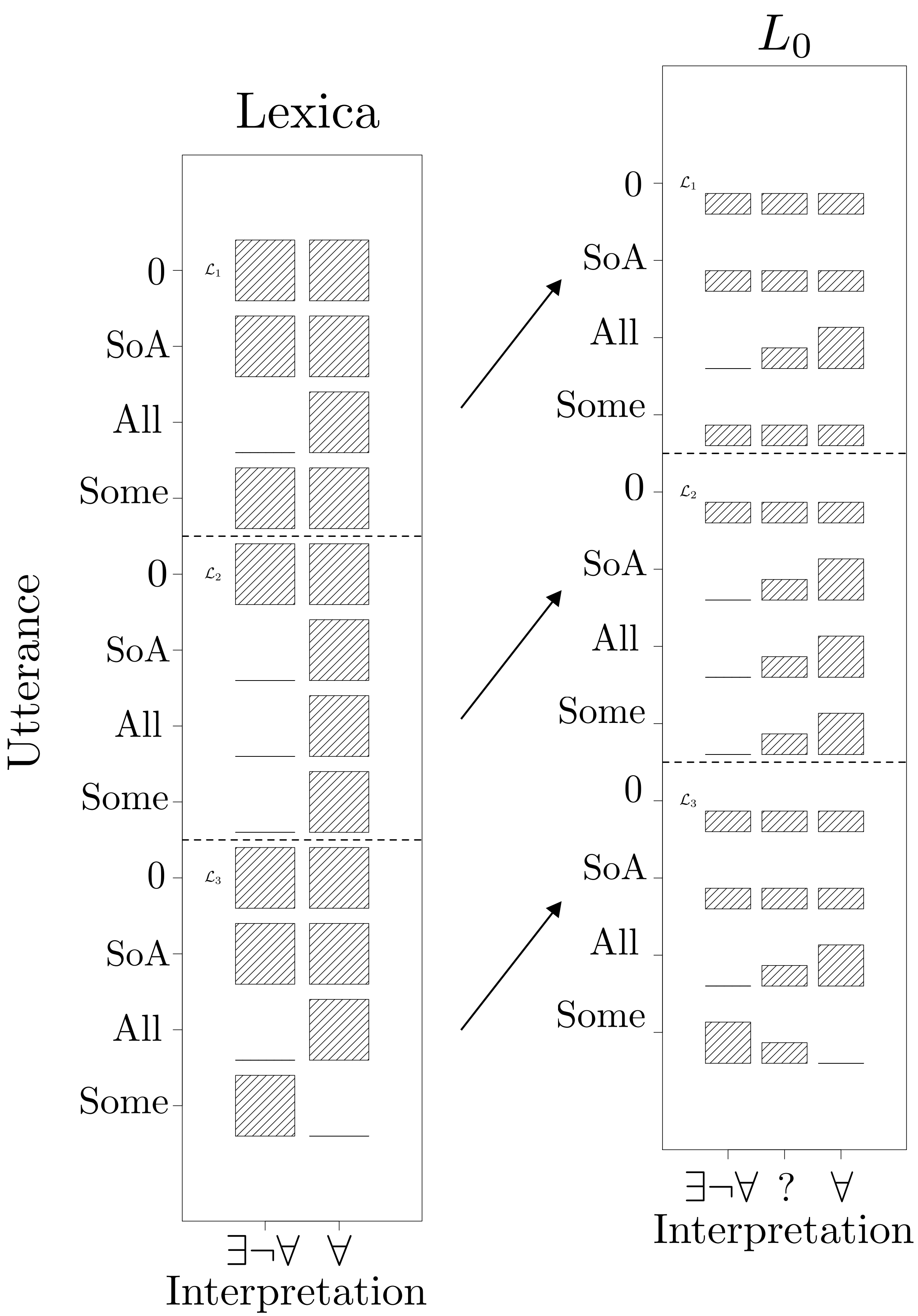
Derivation of embedded implicatures



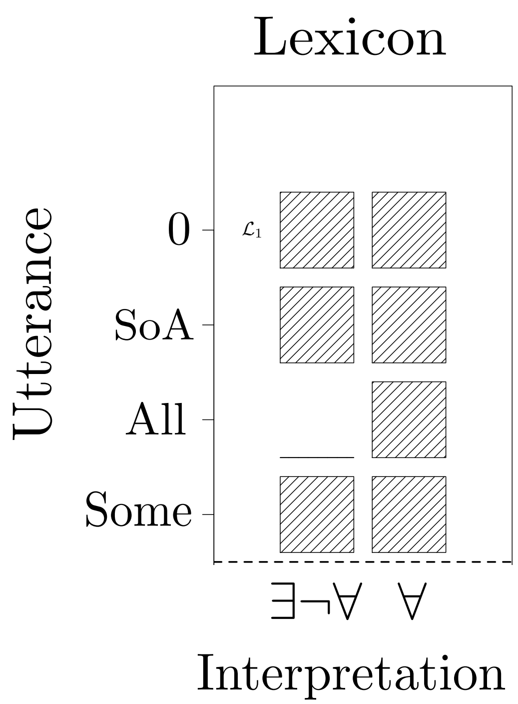
Derivation of embedded implicatures

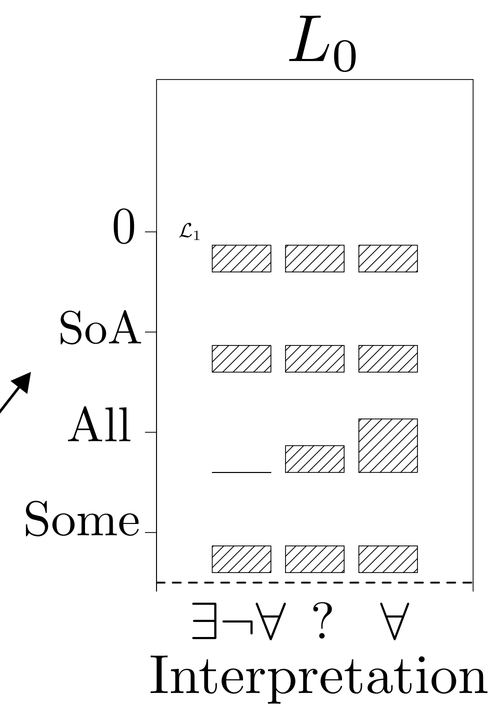


Derivation of embedded implicatures

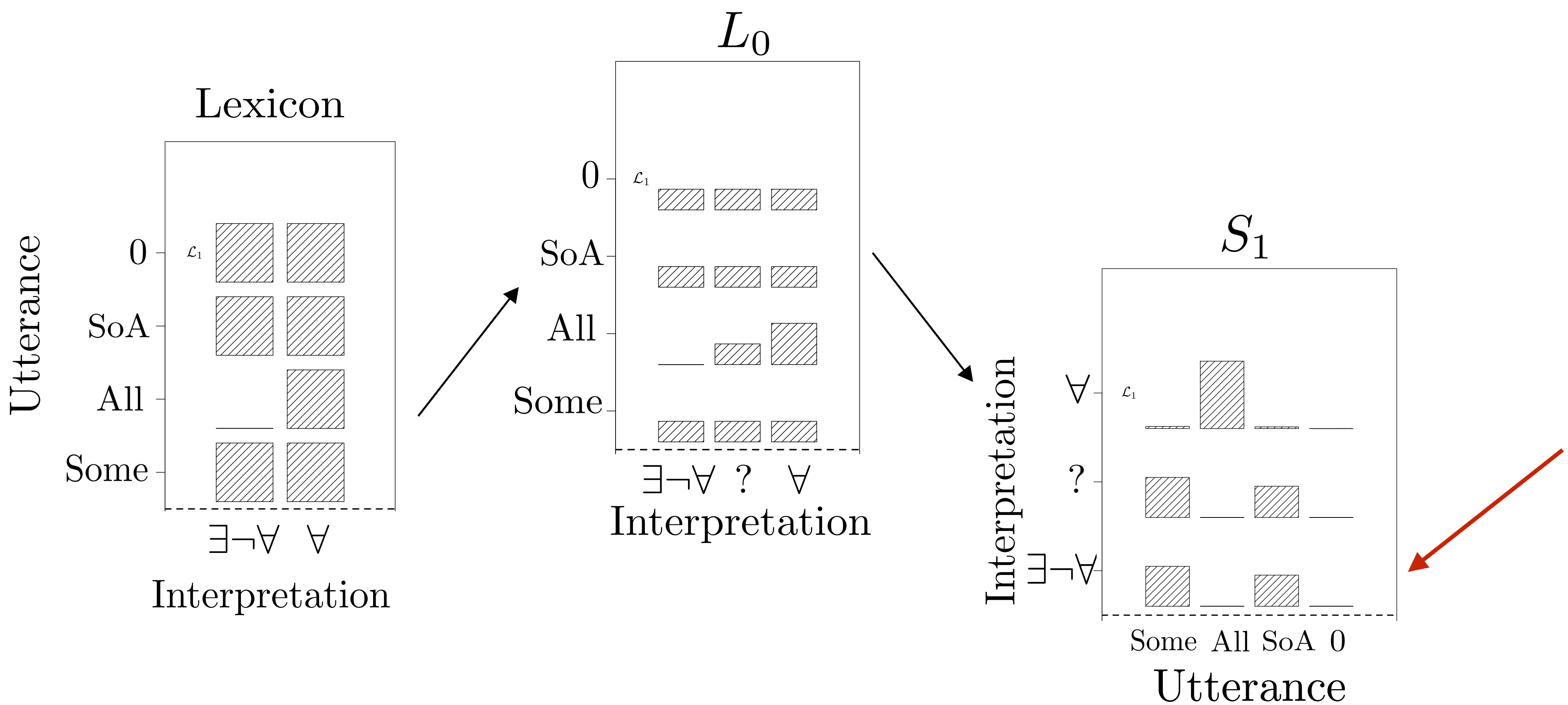


Derivation of embedded implicatures

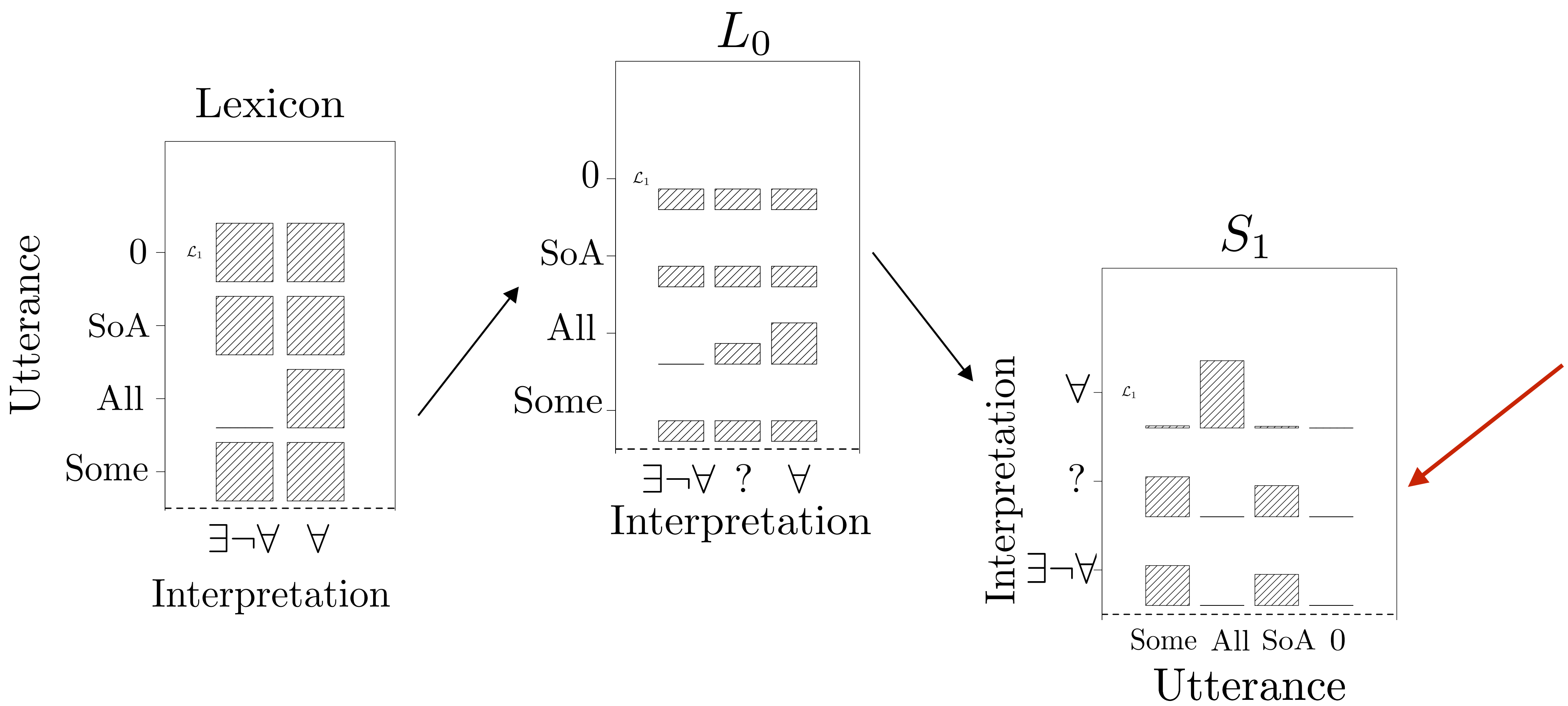




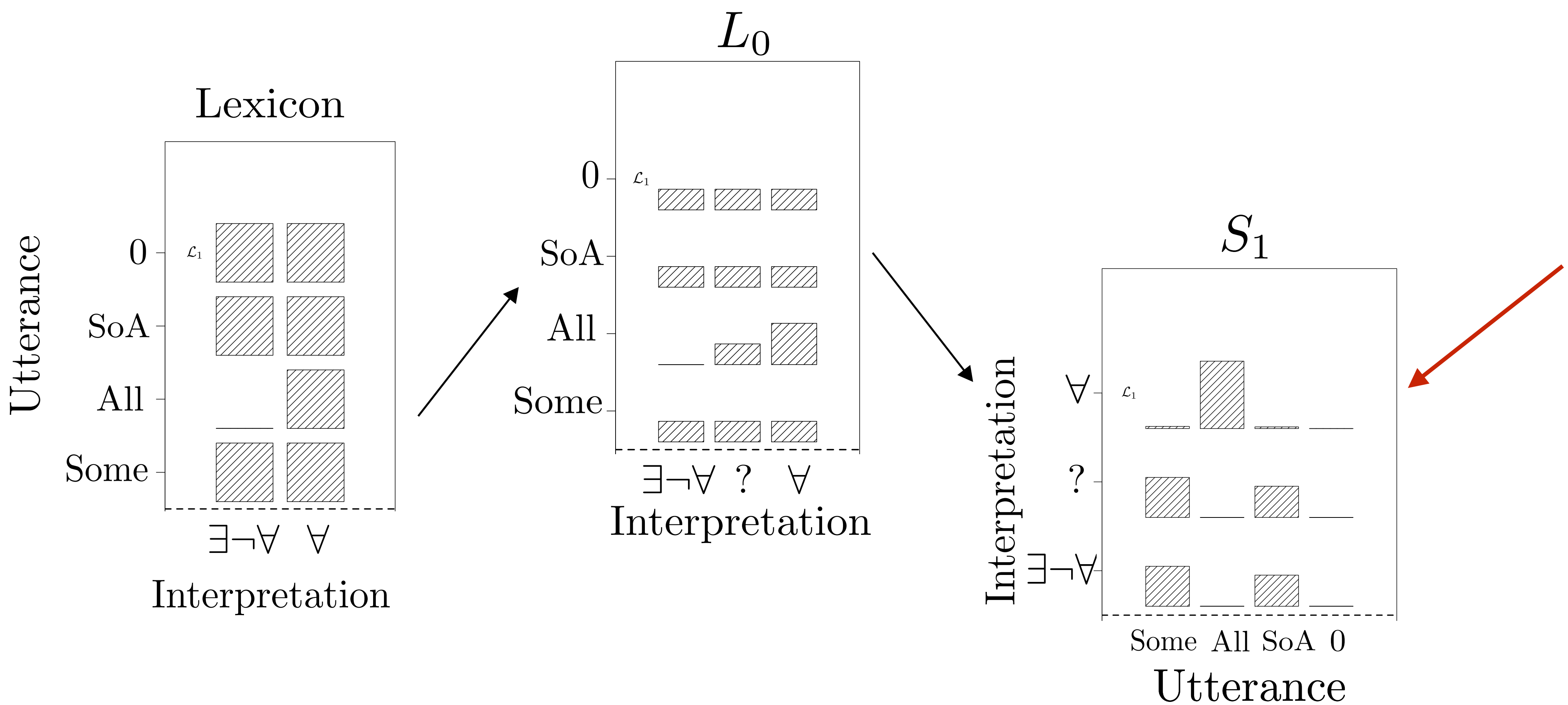
Derivation of embedded implicatures



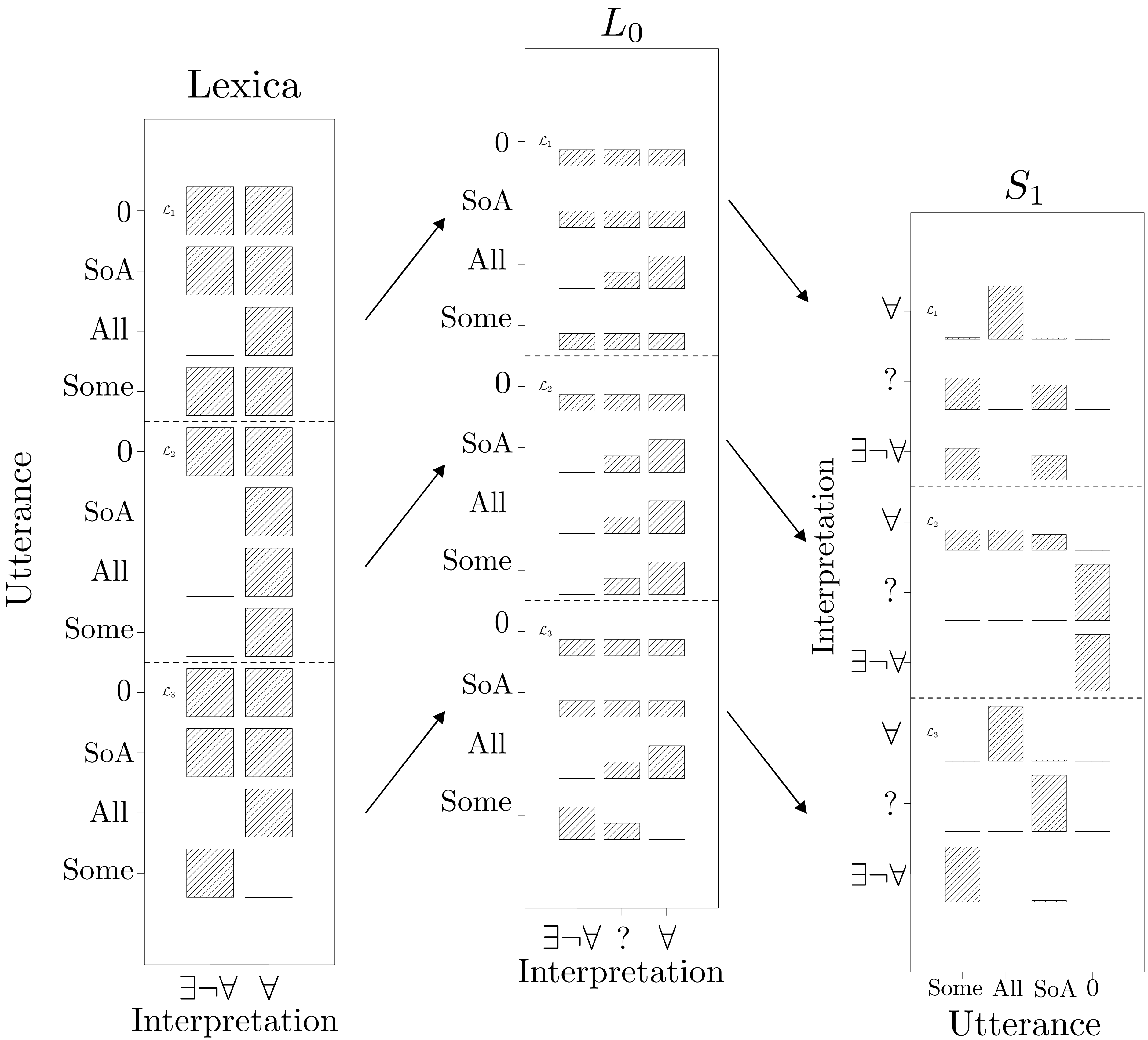
Derivation of embedded implicatures



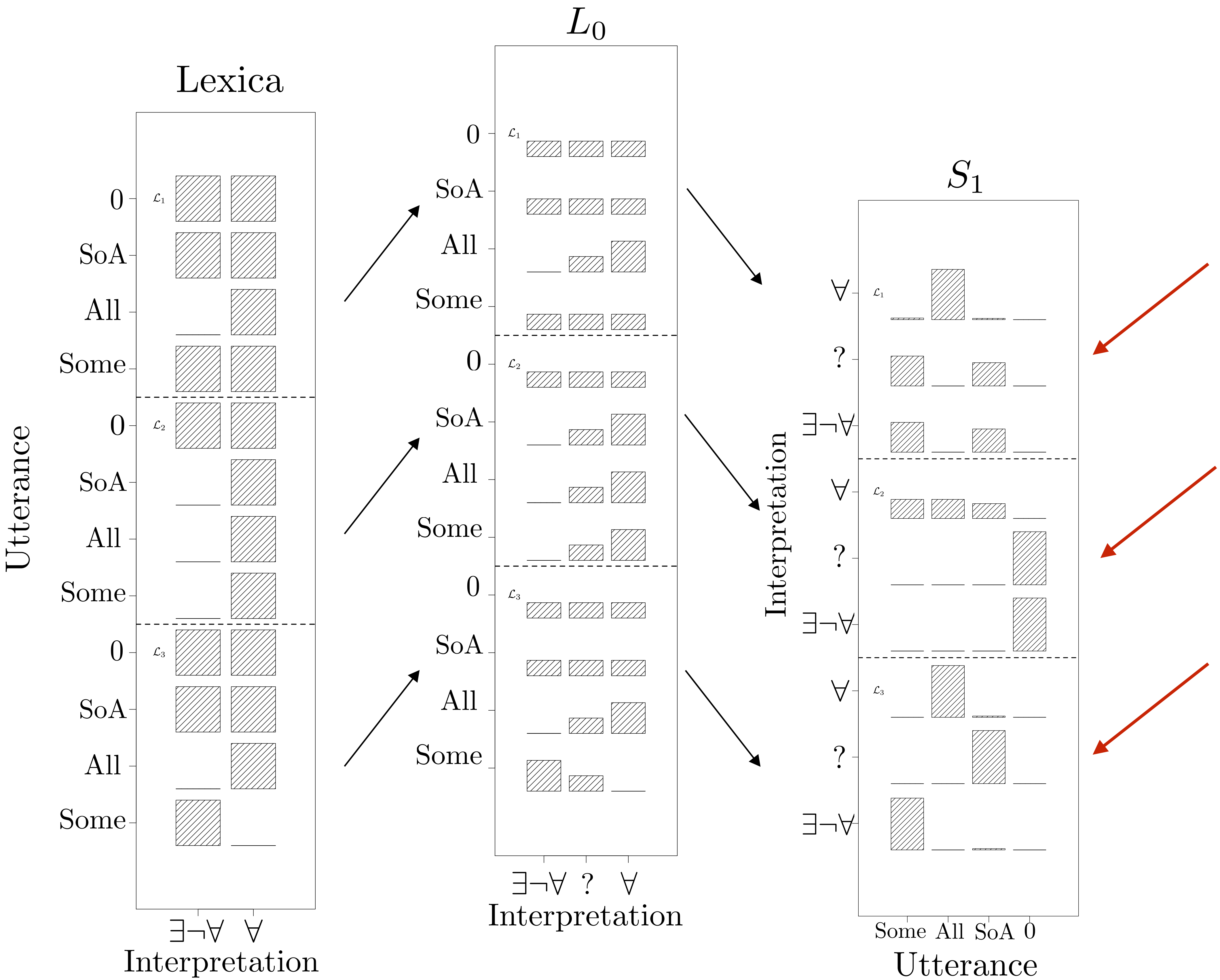
Derivation of embedded implicatures



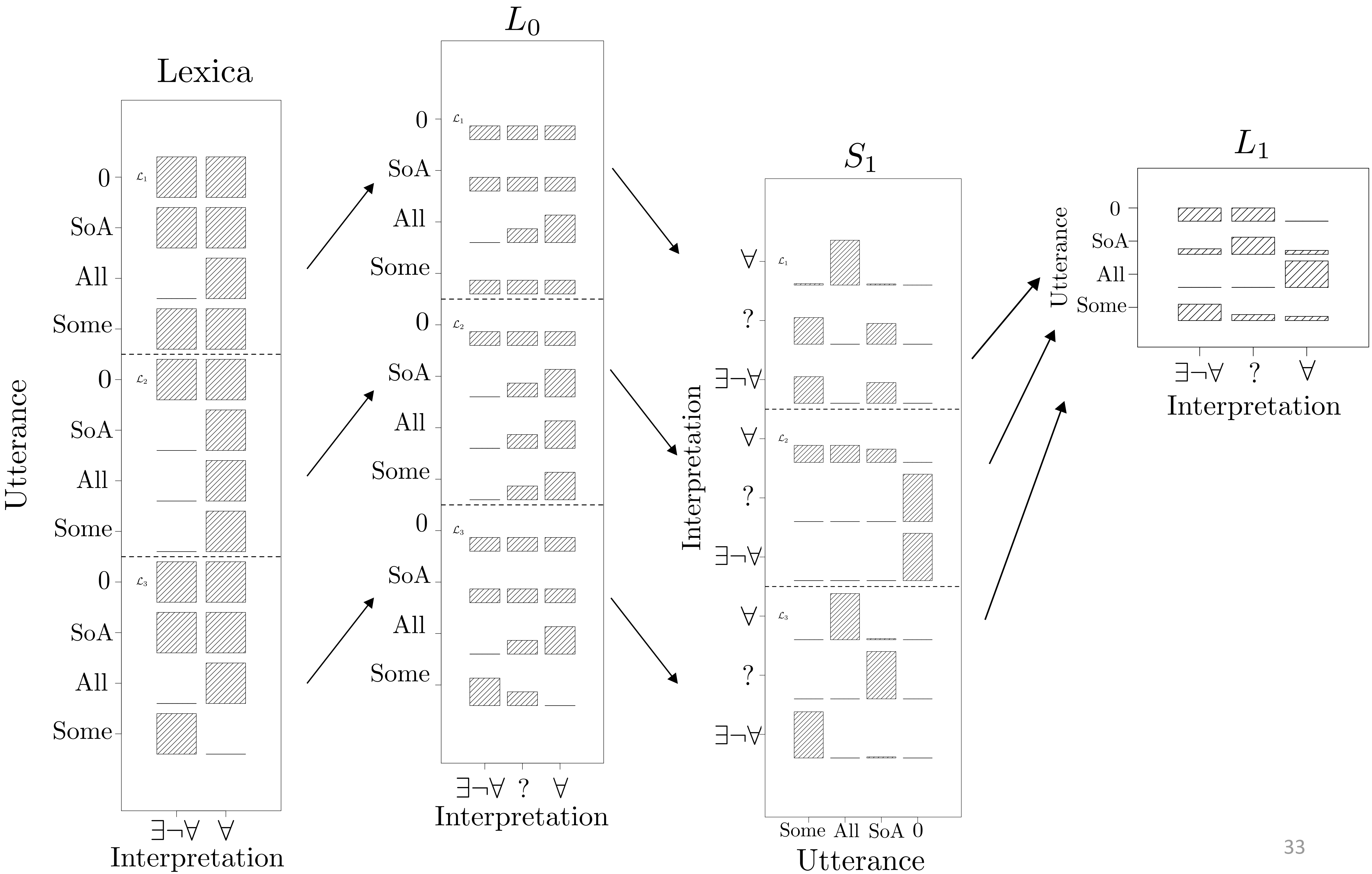
Derivation of embedded implicatures



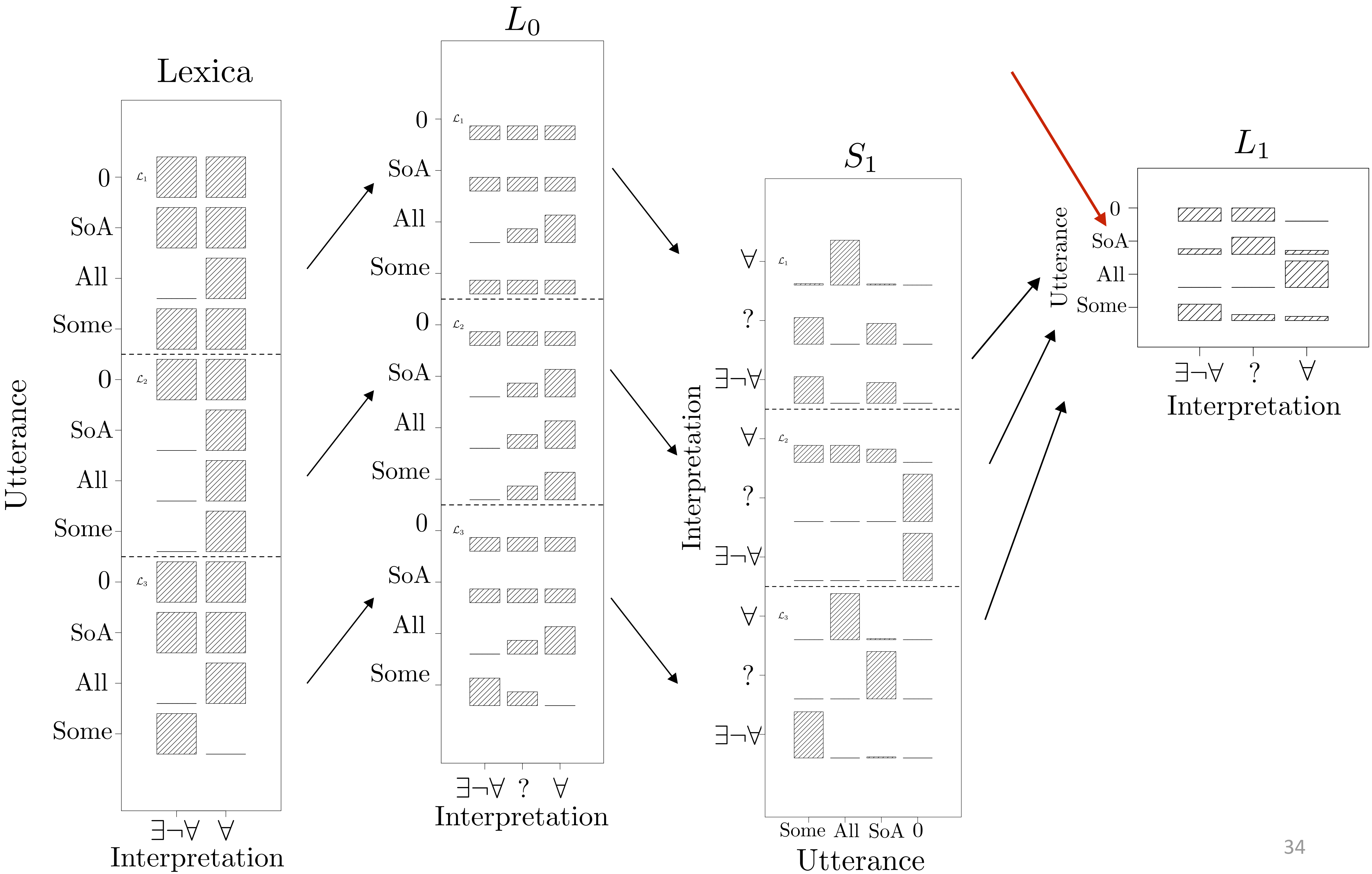
Derivation of embedded implicatures



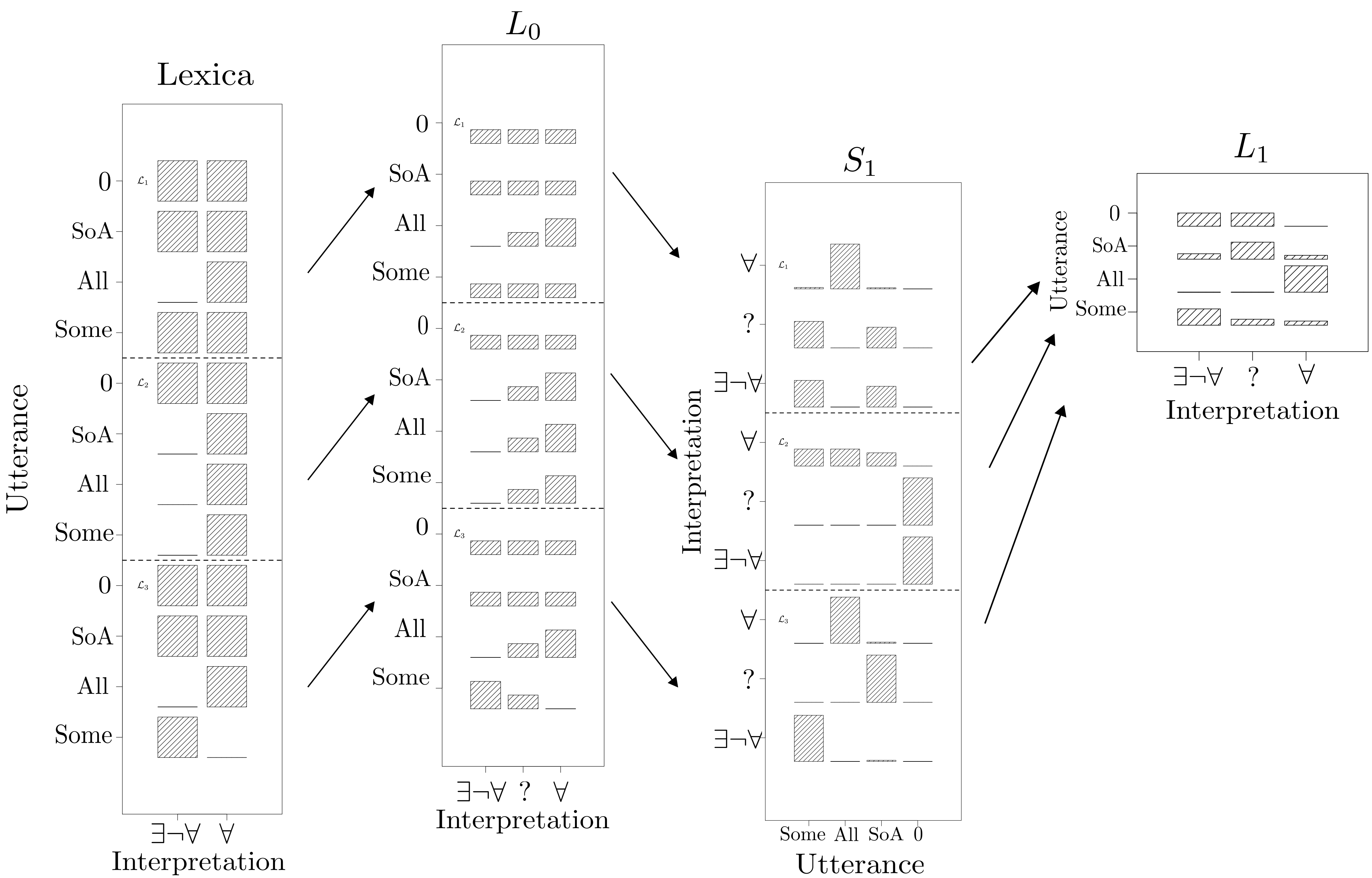
Derivation of embedded implicatures



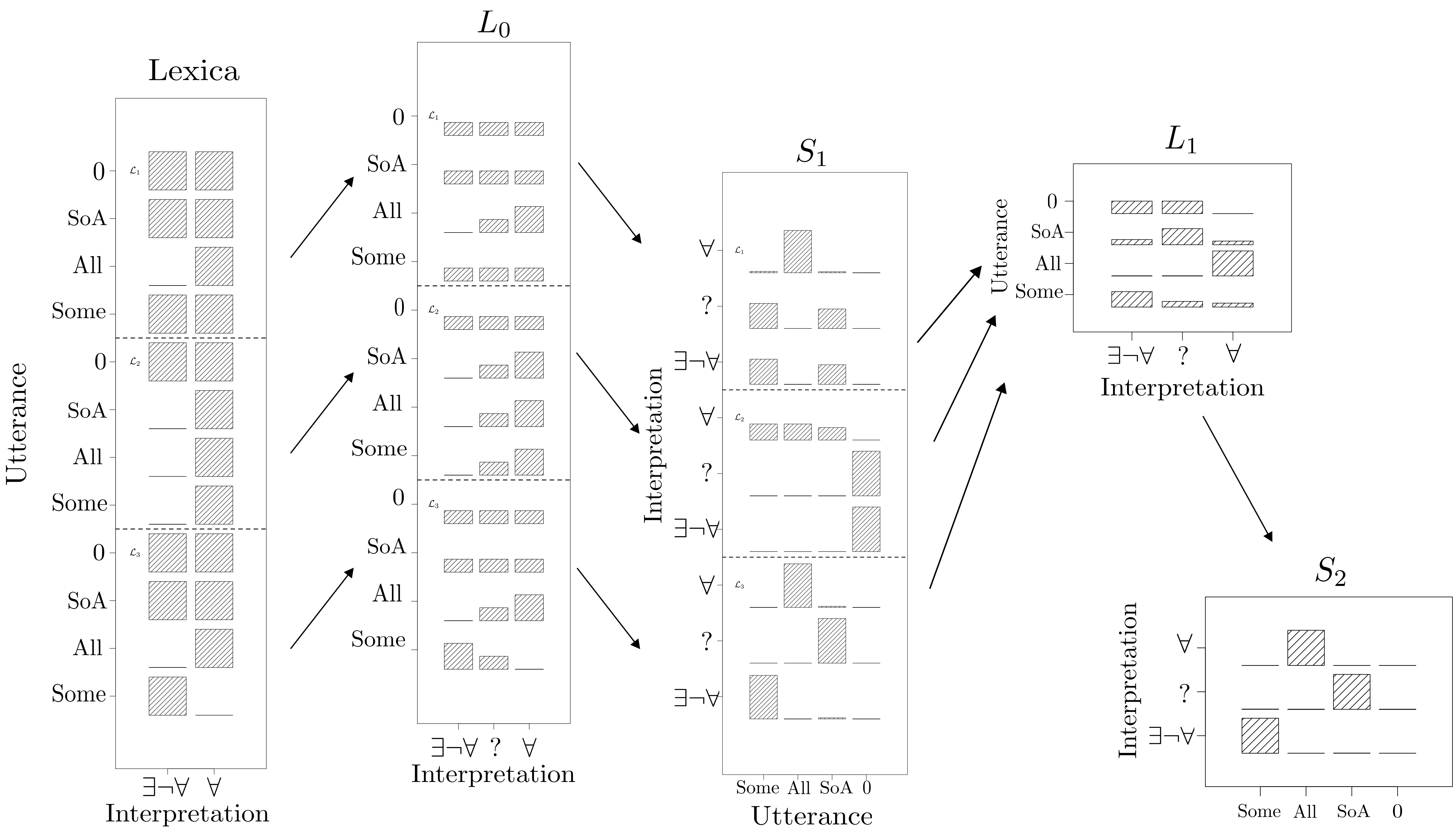
Derivation of embedded implicatures



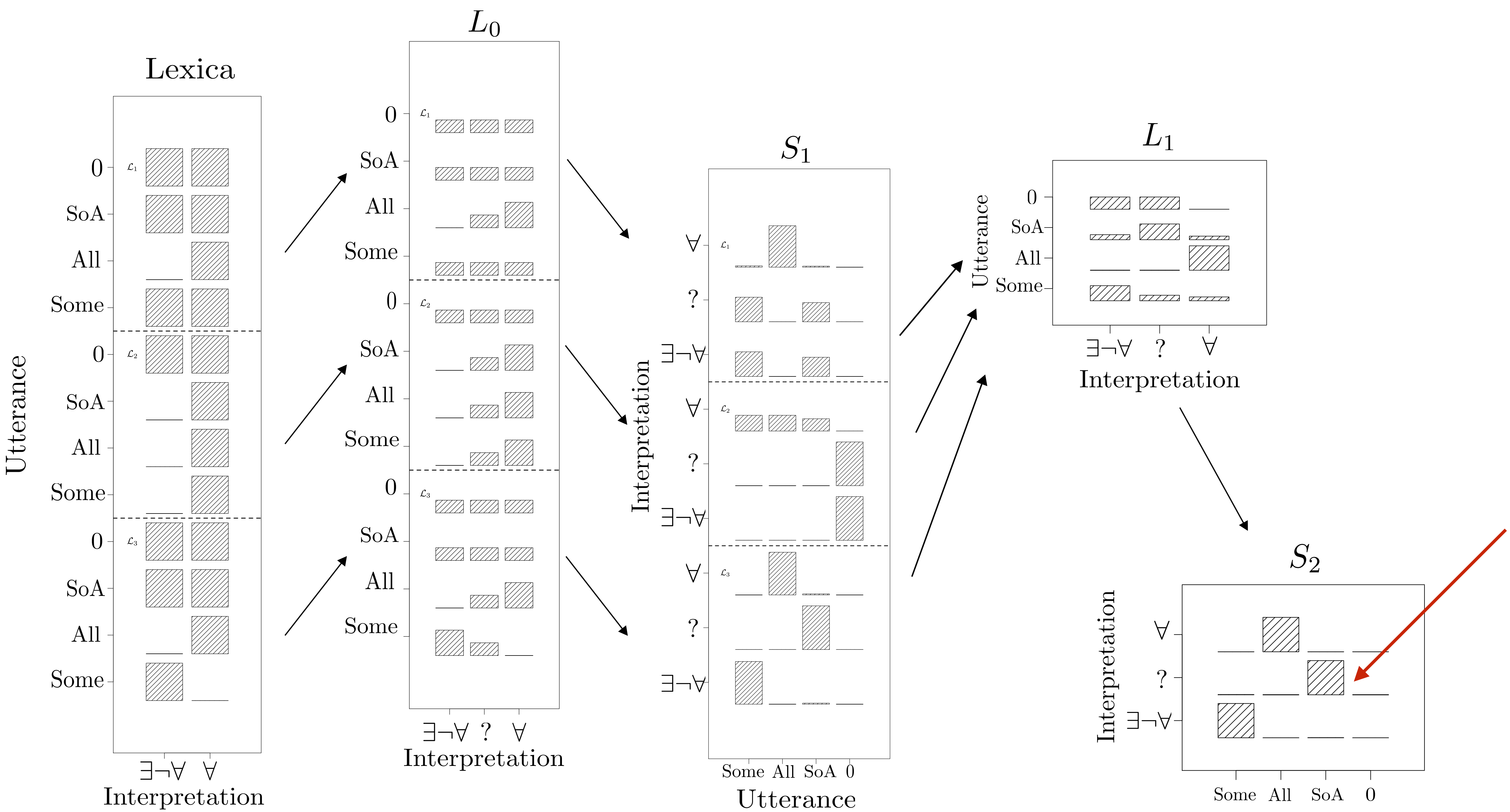
Derivation of embedded implicatures



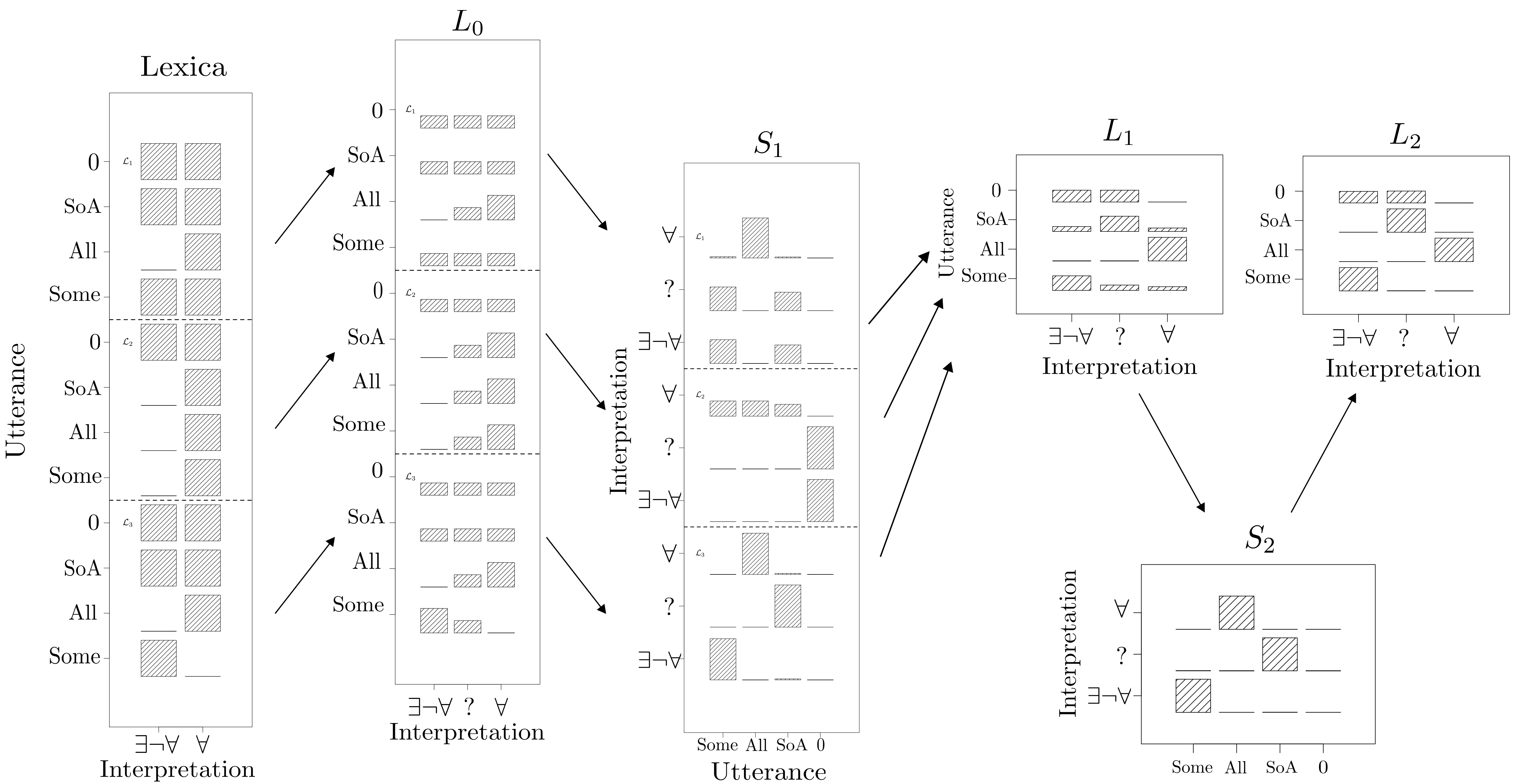
Derivation of embedded implicatures



Derivation of embedded implicatures



Derivation of embedded implicatures



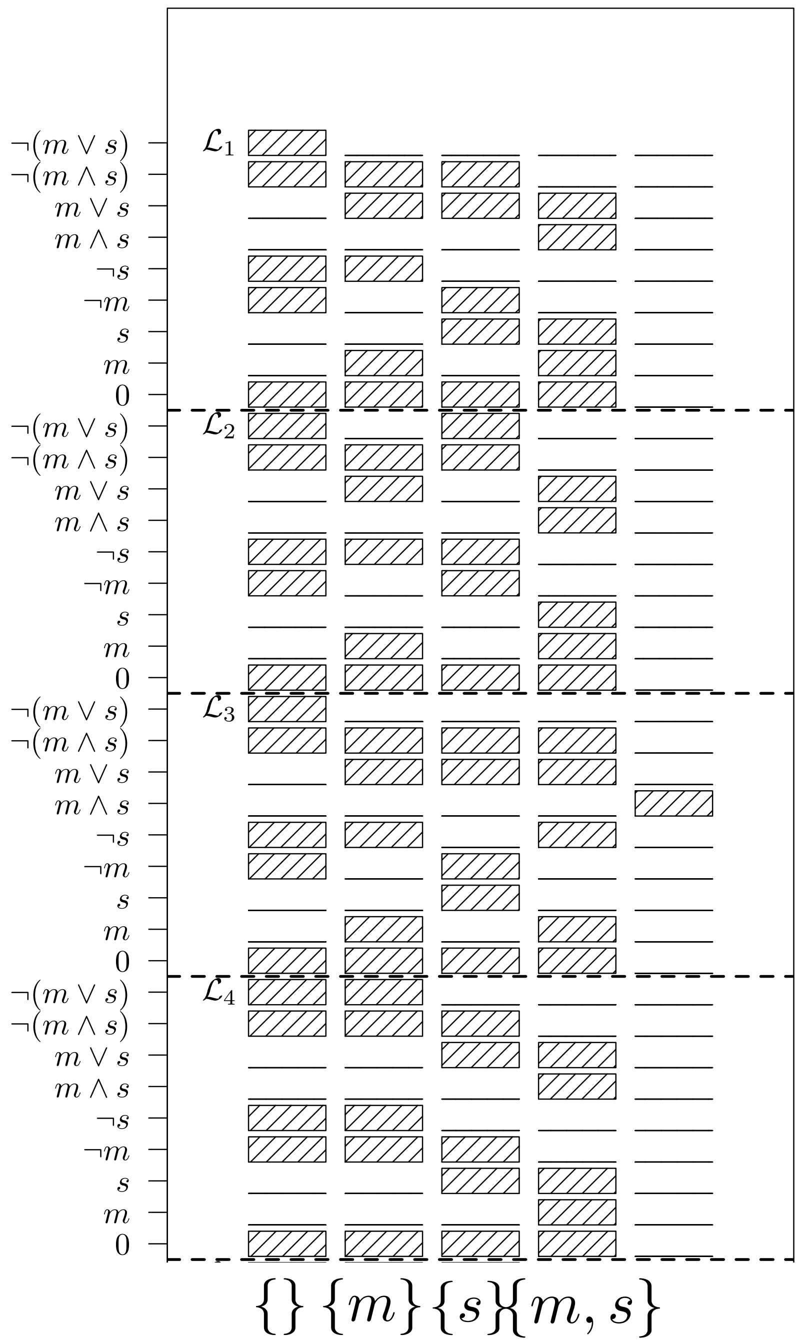
Downward-entailing contexts

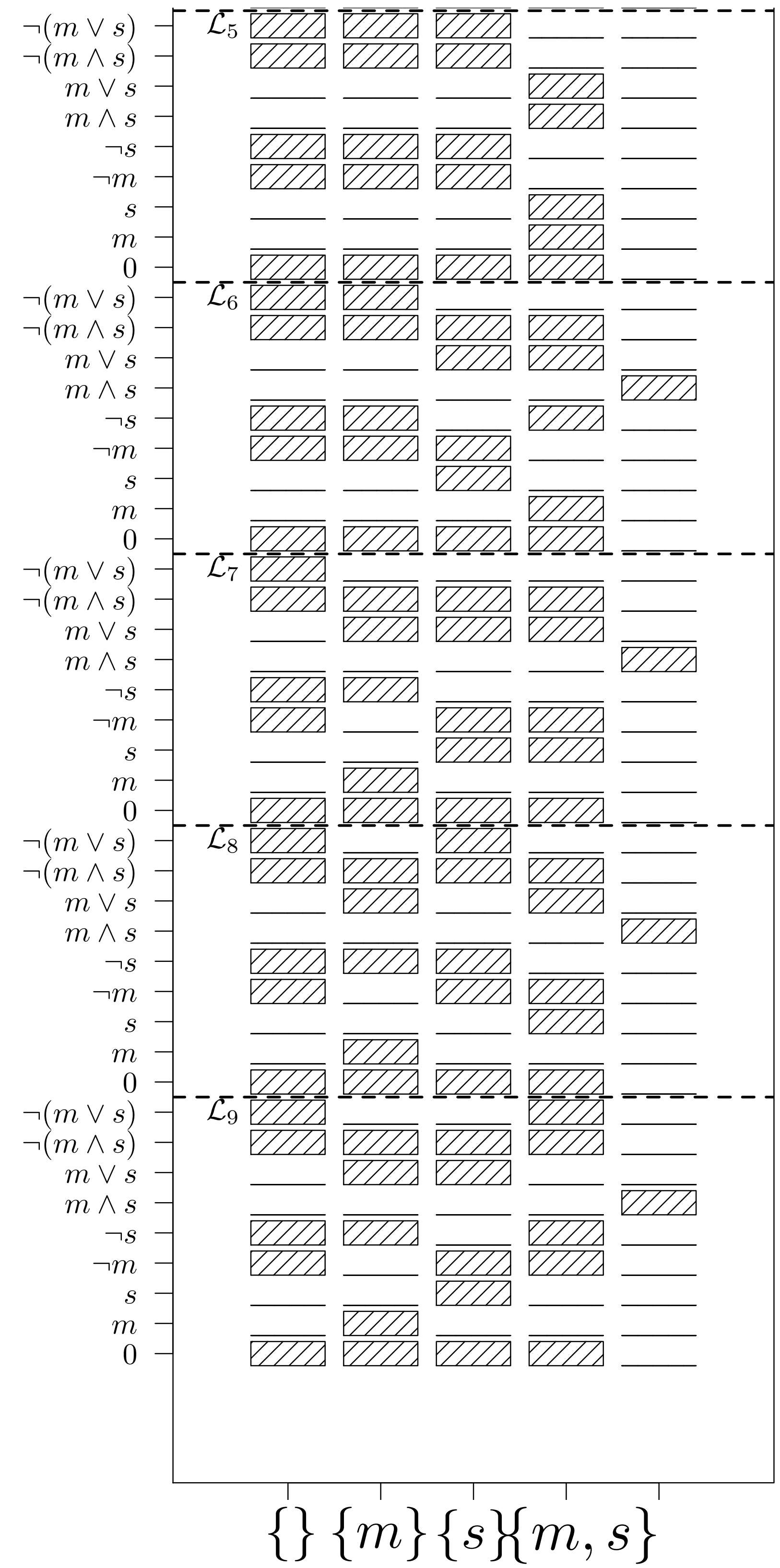
John didn't talk to Mary or Sue.

- Not compatible with John having talked to both Mary and Sue.

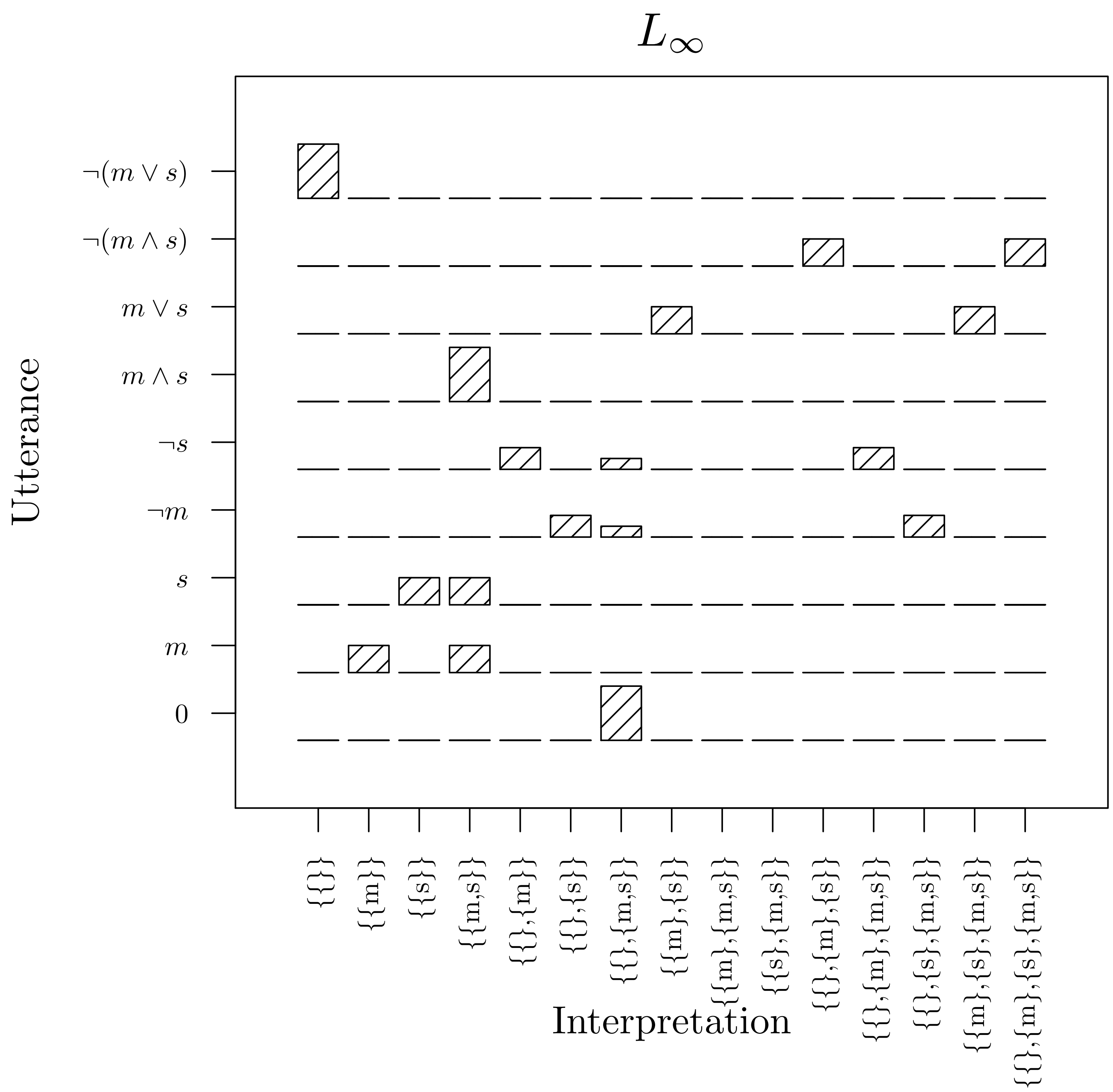
Does the lexical uncertainty framework predict that embedded implicatures will be available in downward-entailing contexts?

Downward-entailing contexts





Downward-entailing contexts

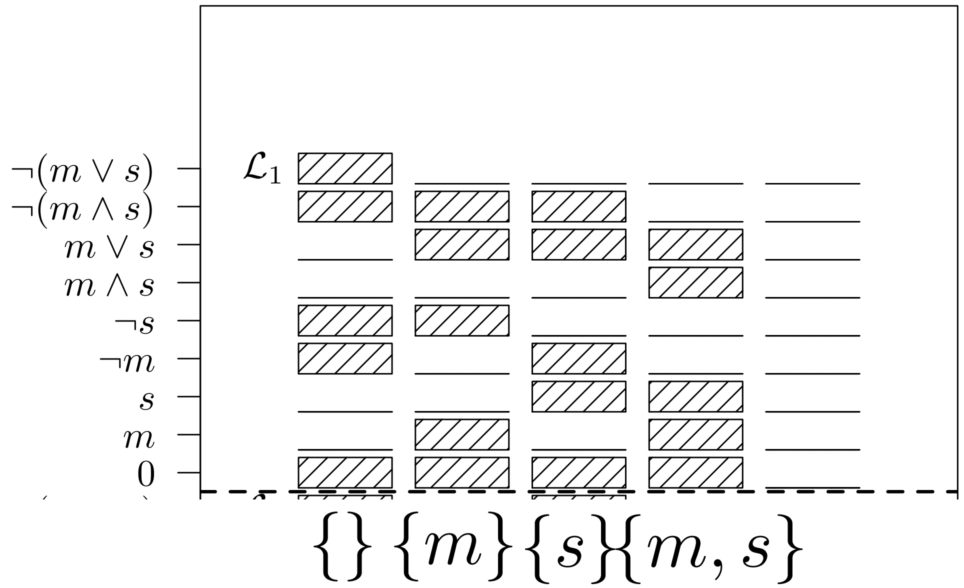


Downward-entailing contexts

Strategy for blocking the DE implicature:

- First show that “not Mary or Sue” will be used to communicate the world in which John talked to neither Mary nor Sue.
- Then show that “not Mary or Sue” will **not** be used to communicate the XOR knowledge state (*either John talked to neither Mary nor Sue, or he talked to both*).

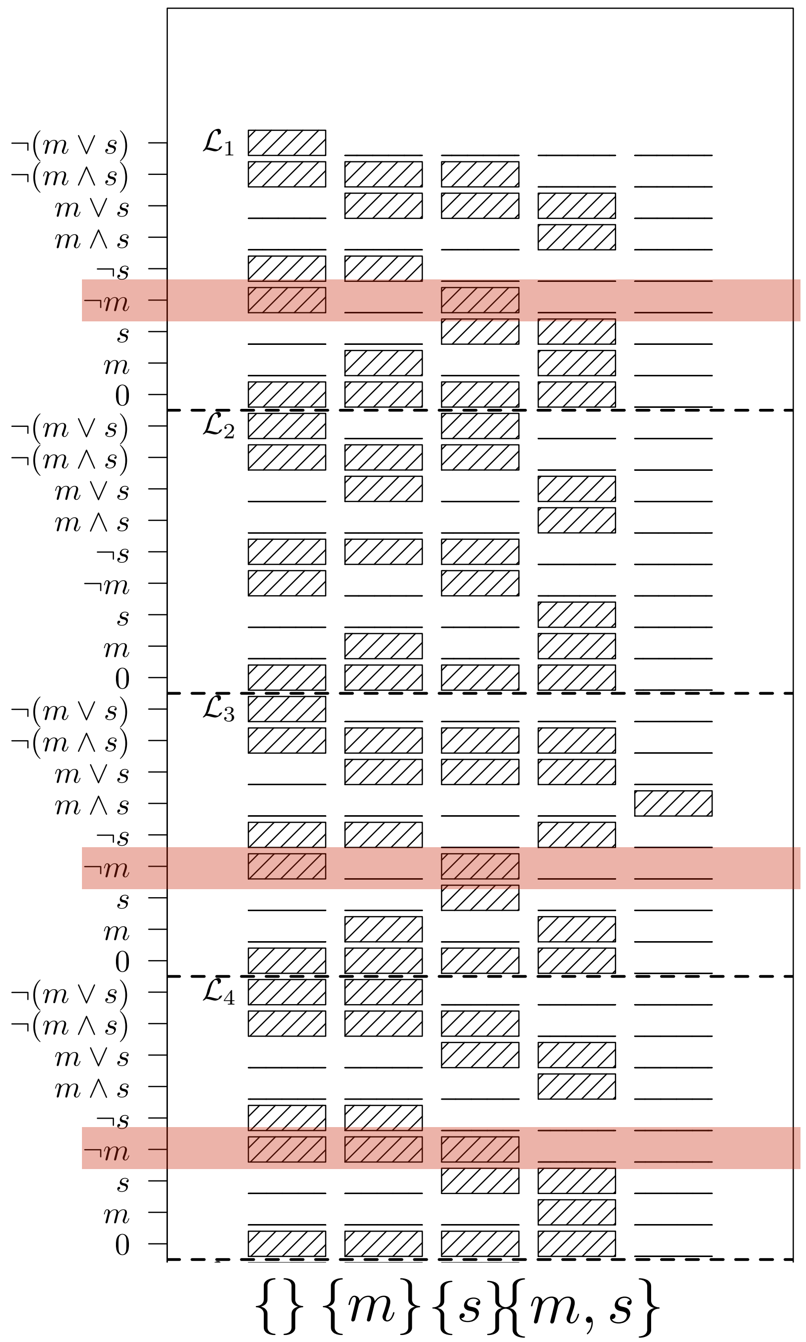
Downward-entailing contexts

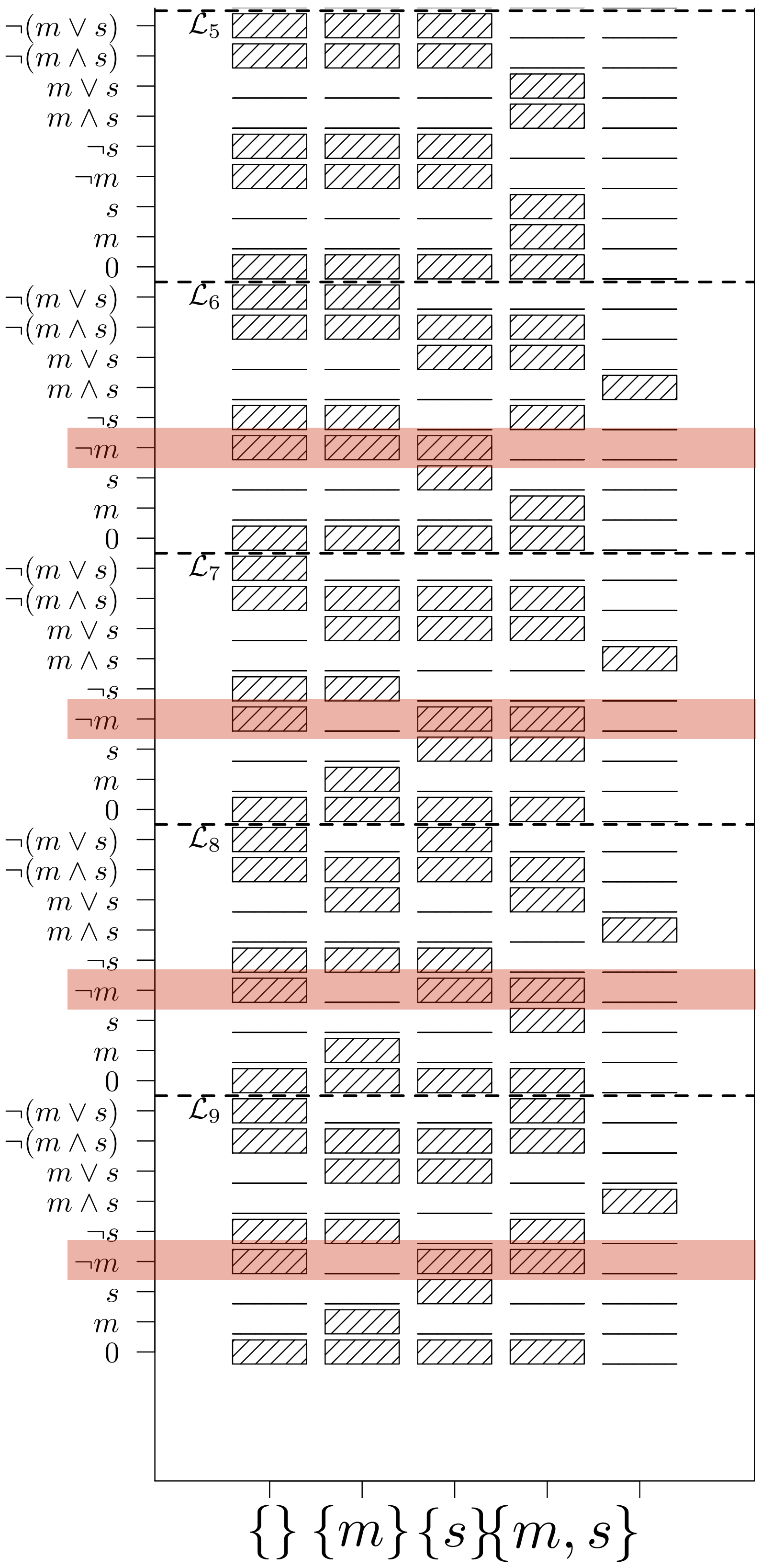


Downward-entailing contexts

“Not Mary or Sue” is *as informative* as “not Mary” in every lexicon, and strictly more informative in most of the lexica.

Downward-entailing contexts

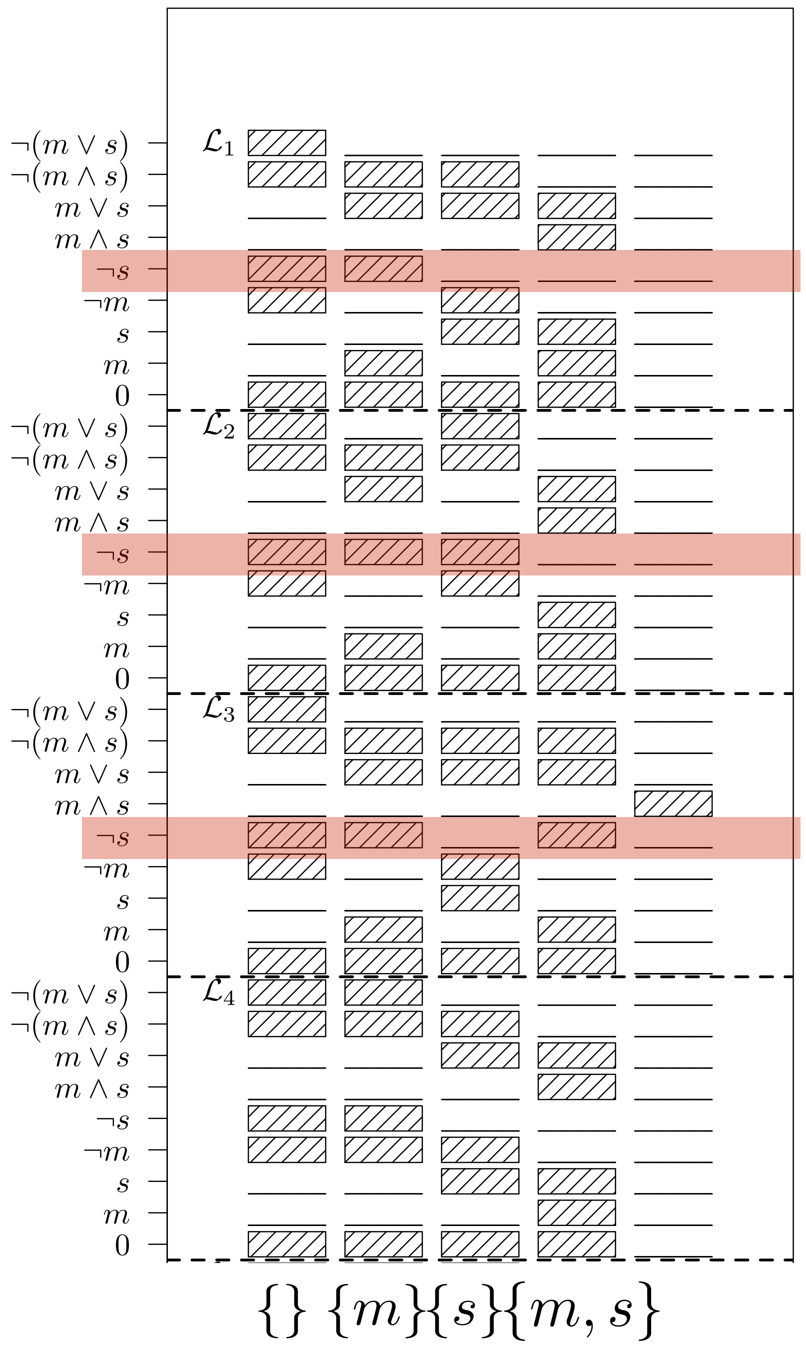


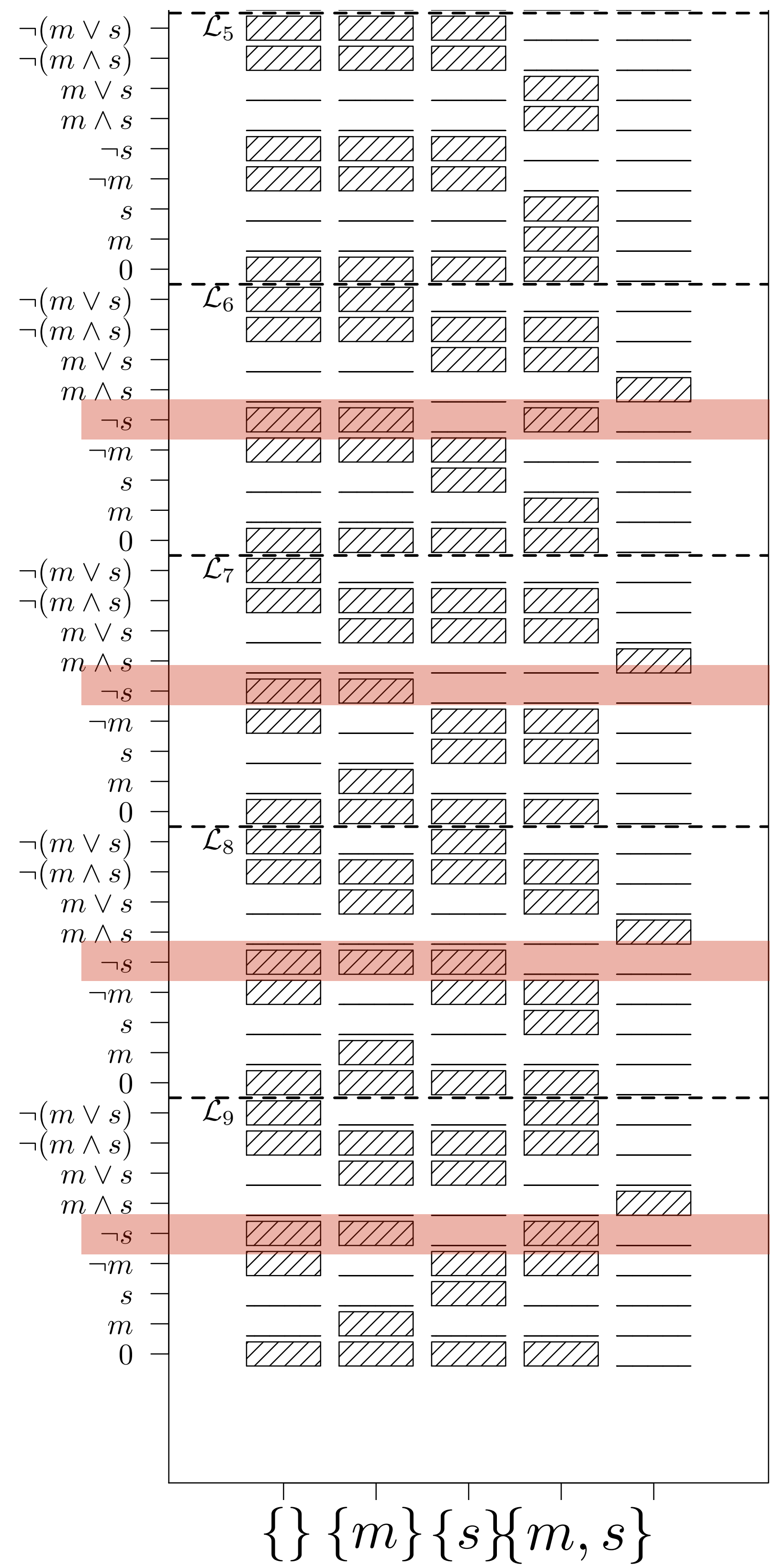


Downward-entailing contexts

"Not Mary or Sue" is *as informative* as "not Sue" in every lexicon, and strictly more informative in most of the lexica.

Downward-entailing contexts





Downward-entailing contexts

We have shown the first point:

- “Not Mary or Sue” will be used to communicate the world in which John talked to neither Mary nor Sue.

Downward-entailing contexts

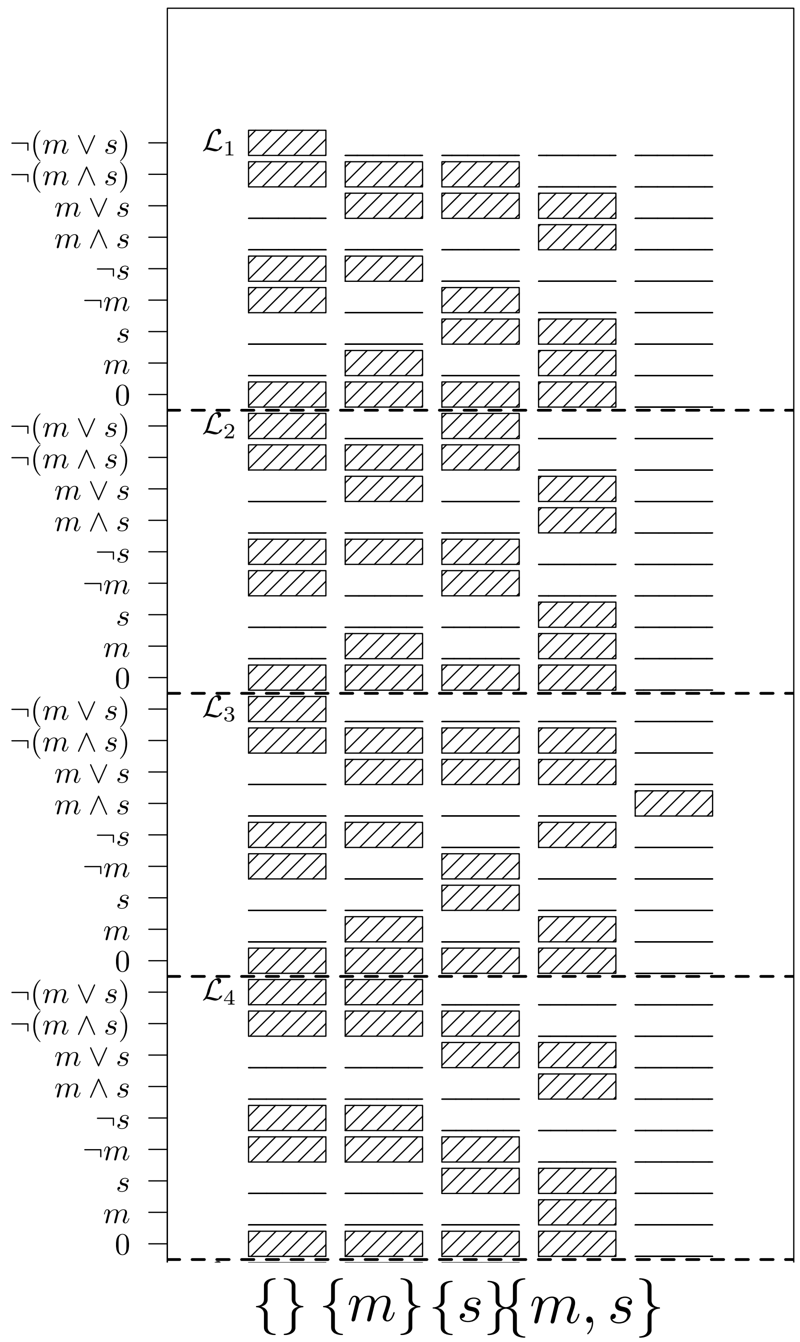
Now we show the second point:

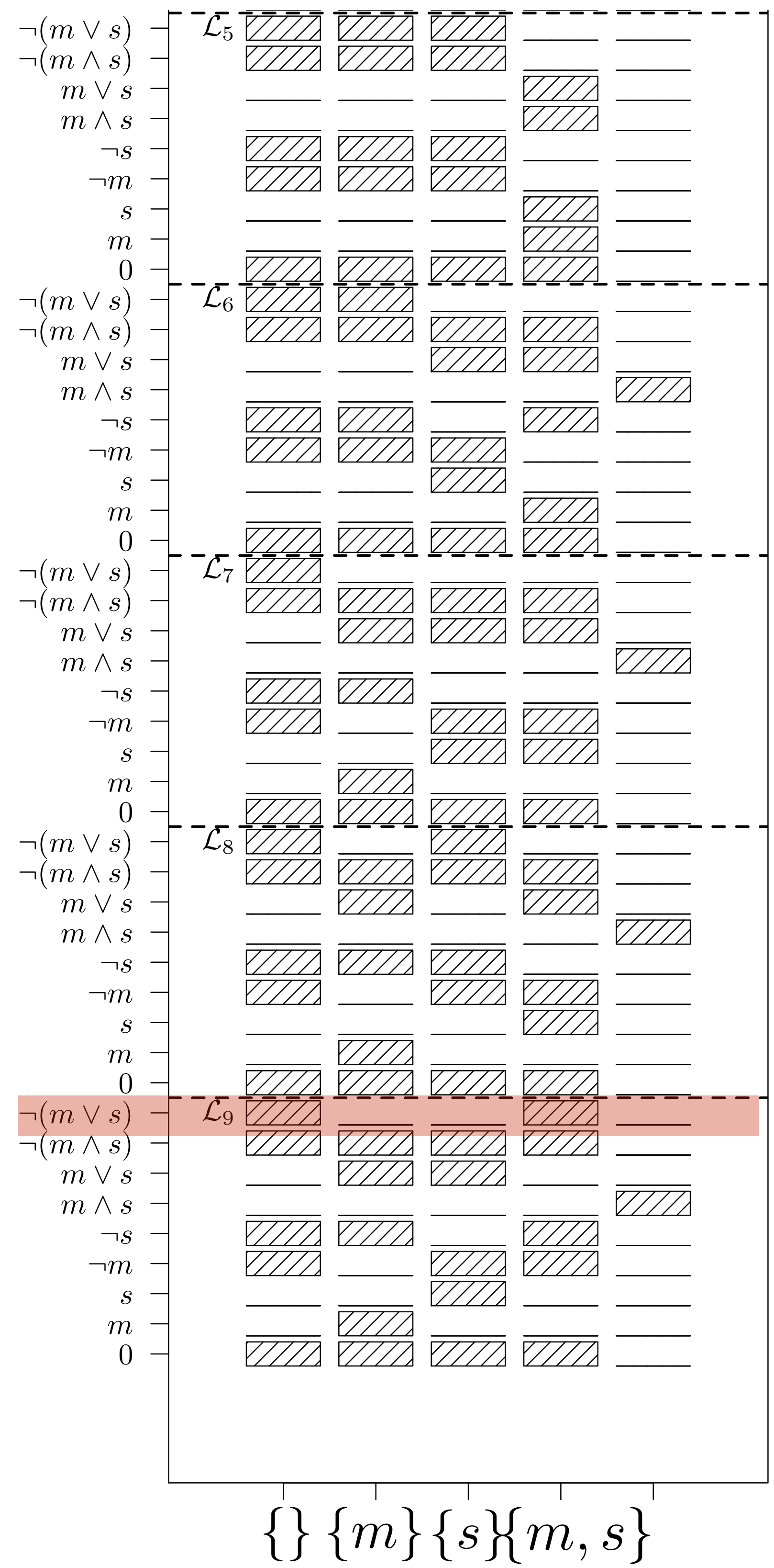
- “Not Mary or Sue” will **not** be used to communicate the XOR knowledge state (*either John talked to neither Mary nor Sue, or he talked to both*).

Downward-entailing contexts

There is only one lexicon in which "not Mary or Sue" is compatible with the XOR knowledge state (*either John talked to neither Mary nor Sue, or John talked to both Mary and Sue*).

Downward-entailing contexts

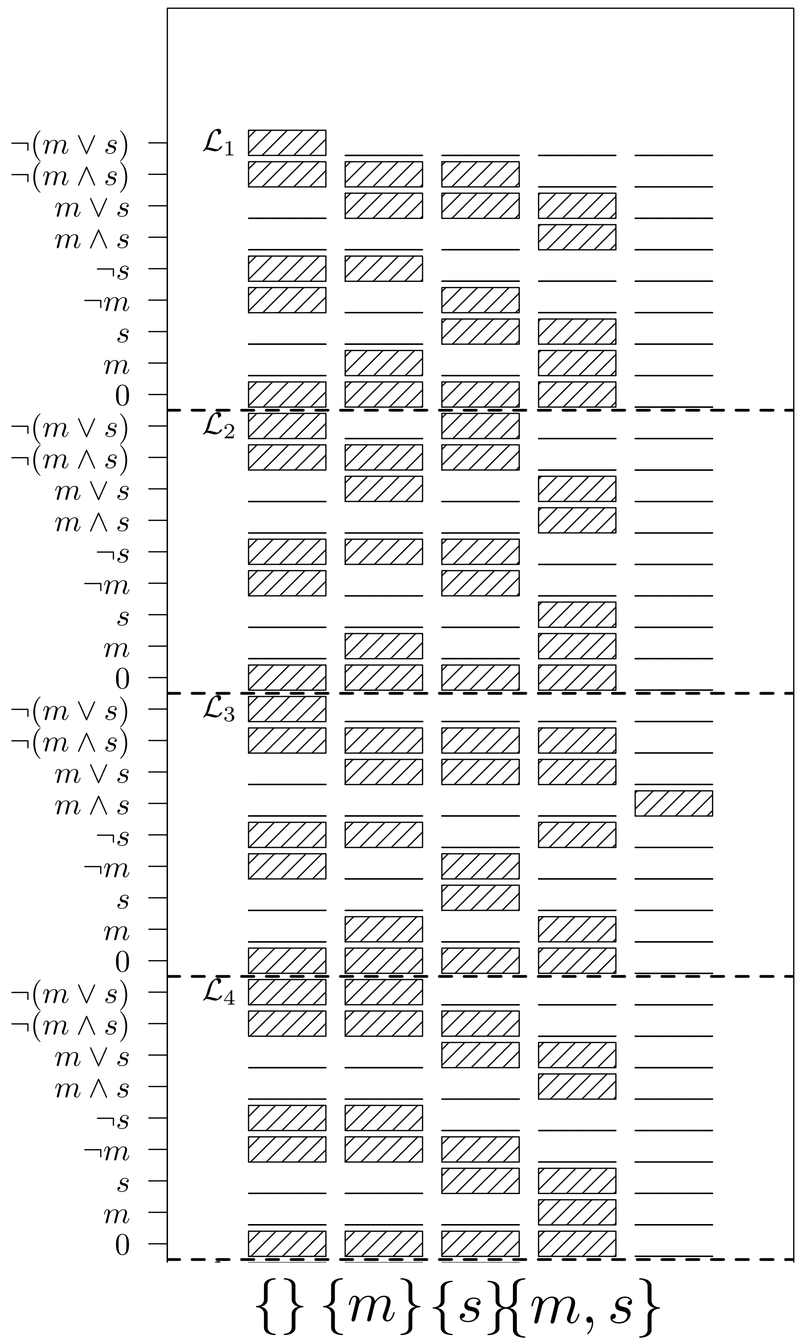


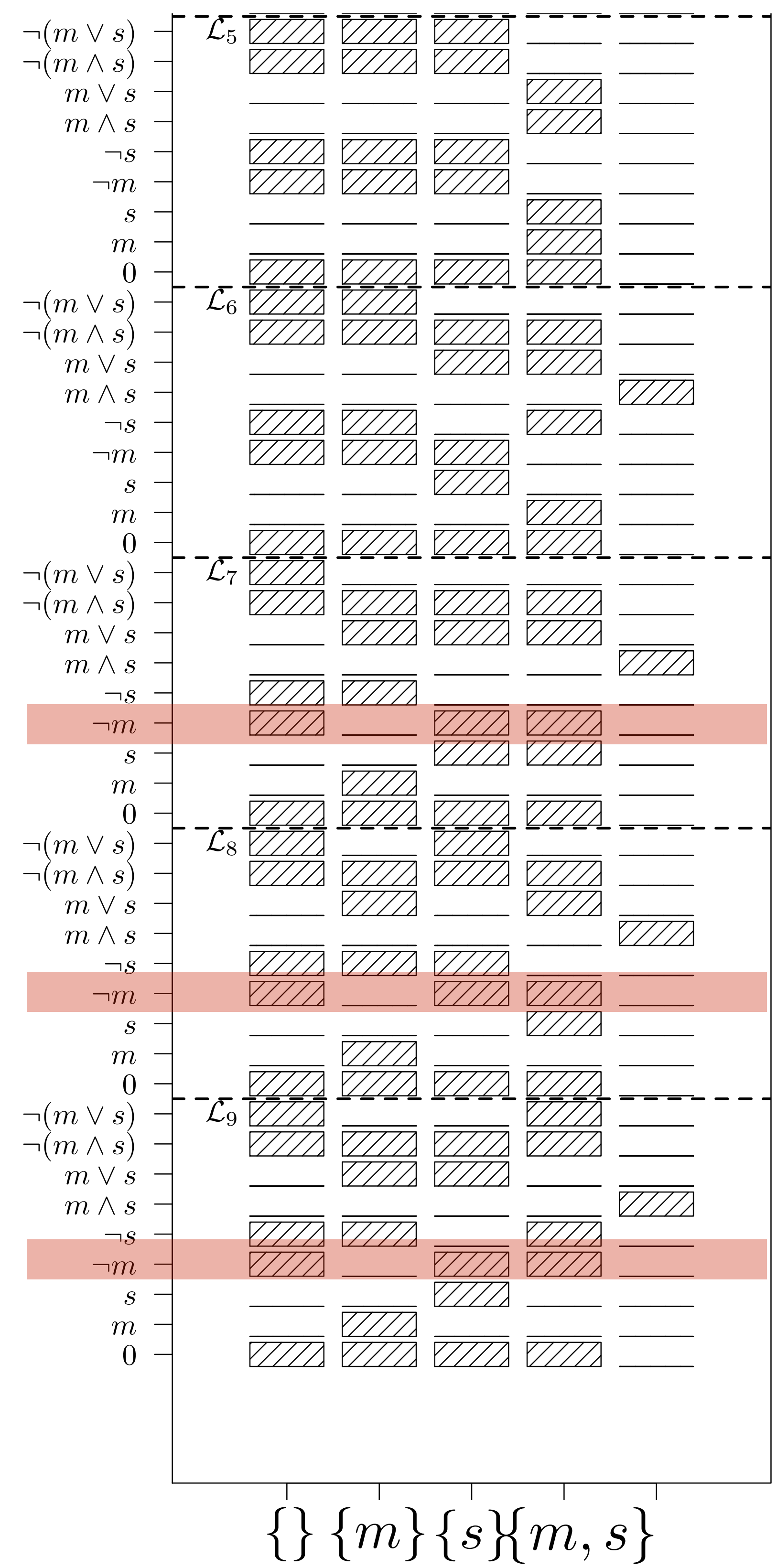


Downward-entailing contexts

There are three lexica in which “not Mary” is compatible with the XOR knowledge state (*either John talked to neither Mary nor Sue, or John talked to both Mary and Sue*).

Downward-entailing contexts

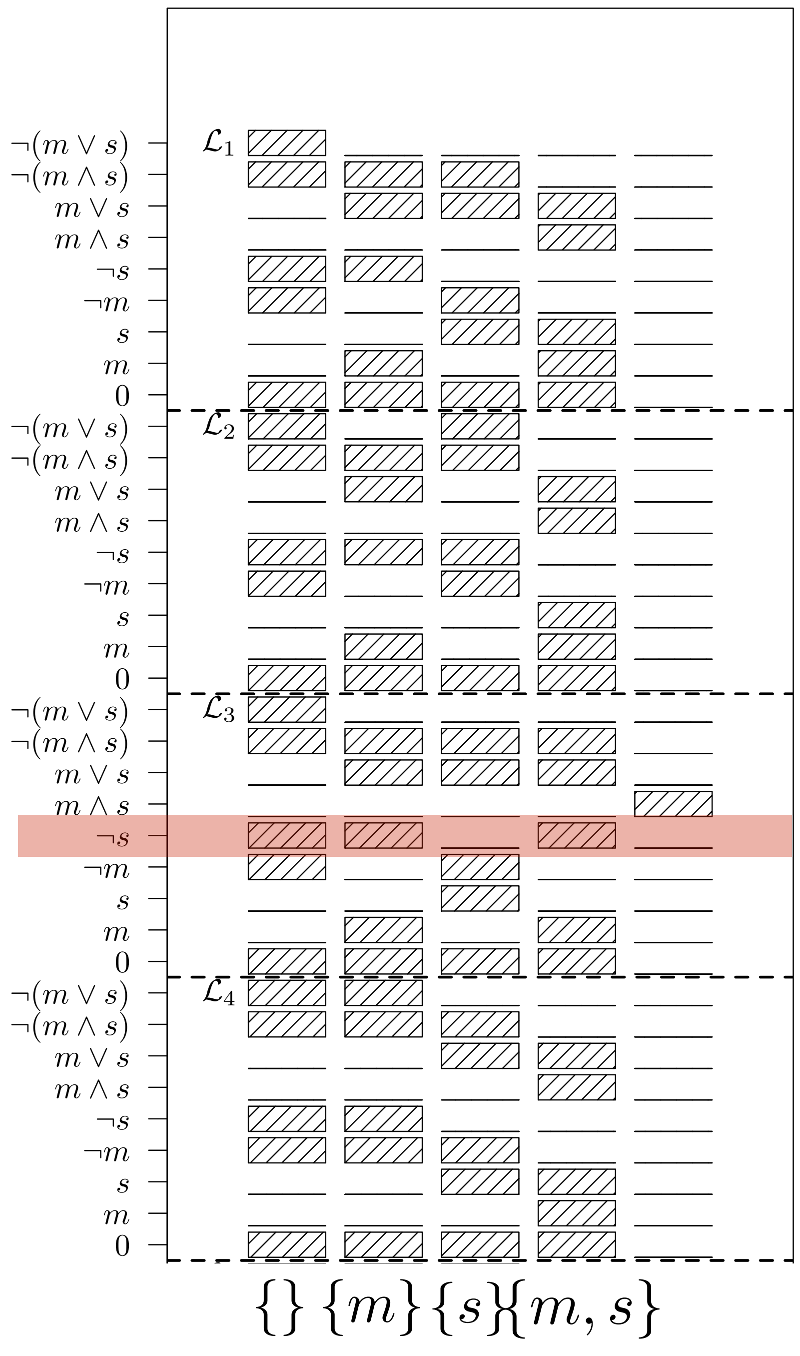


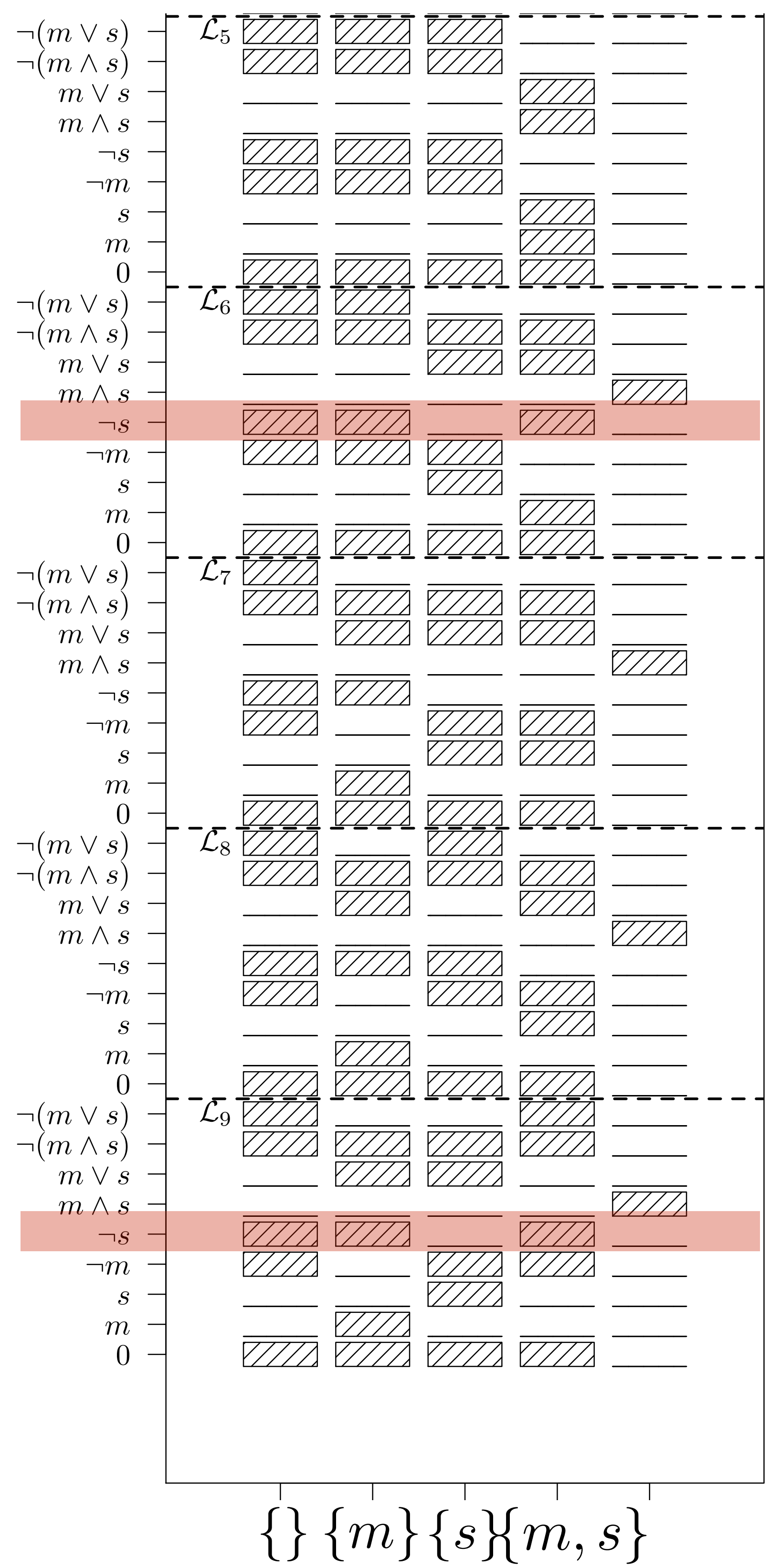


Downward-entailing contexts

There are three lexica in which “not Sue” is compatible with the XOR knowledge state (*either John talked to neither Mary nor Sue, or John talked to both Mary and Sue*).

Downward-entailing contexts





Downward-entailing contexts

We have now demonstrated both points:

- “Not Mary or Sue” will be used to communicate the world in which John talked to neither Mary nor Sue.
- “Not Mary or Sue” will **not** be used to communicate the XOR knowledge state (*either John talked to neither Mary nor Sue, or he talked to both*).

Potts et al (2015)

L

	N	S	A
B scored	0	.71	.29
B aced	0	0	1
0	.75	.25	0

s_1

	B scored	*B aced*	**0**
N	0	0	1
S	1	0	0
A	.33	.67	0

	B scored	*B aced*	**0**
N	0	0	1
S	1	0	0
A	0	1	0

	B scored	*B aced*	**0**
N	0	0	1
S	0	0	1
A	.5	.5	0

l_0

	N	S	A
B scored	0	.5	.5
B aced	0	0	1
0	.33	.33	.33

	N	S	A
B scored	0	1	0
B aced	0	0	1
0	.33	.33	.33

	N	S	A
B scored	0	0	1
B aced	0	0	1
0	.33	.33	.33

M

	N	S	A
B scored	F	T	T
B aced	F	F	T
0	T	T	T

	N	S	A
B scored	F	T	F
B aced	F	F	T
0	T	T	T

	N	S	A
B scored	F	F	T
B aced	F	F	T
0	T	T	T

$\mathcal{L}$

$[\![scored]\!] = \{\langle S, b\rangle, \langle A, b\rangle\}$
$[\![aced]\!] = \{\langle A, b\rangle\}$
$[\![Player\ B]\!]$ = as in tab. 1

$[\![scored]\!] = \{\langle S, b\rangle\}$
$[\![aced]\!] = \{\langle A, b\rangle\}$
$[\![Player\ B]\!]$ = as in tab. 1

$[\![scored]\!] = \{\langle A, b\rangle\}$
$[\![aced]\!] = \{\langle A, b\rangle\}$
$[\![Player\ B]\!]$ = as in tab. 1

Figure 2: Simple scalar inference. We assume a flat prior over states and lexica. $C(\mathbf{0}) = 5$, and $C(m) = 0$ for the other messages. The uncertainty listener L infers that the general term *scored* excludes its specific counterpart *aced* in this context.

Potts et al (2015)

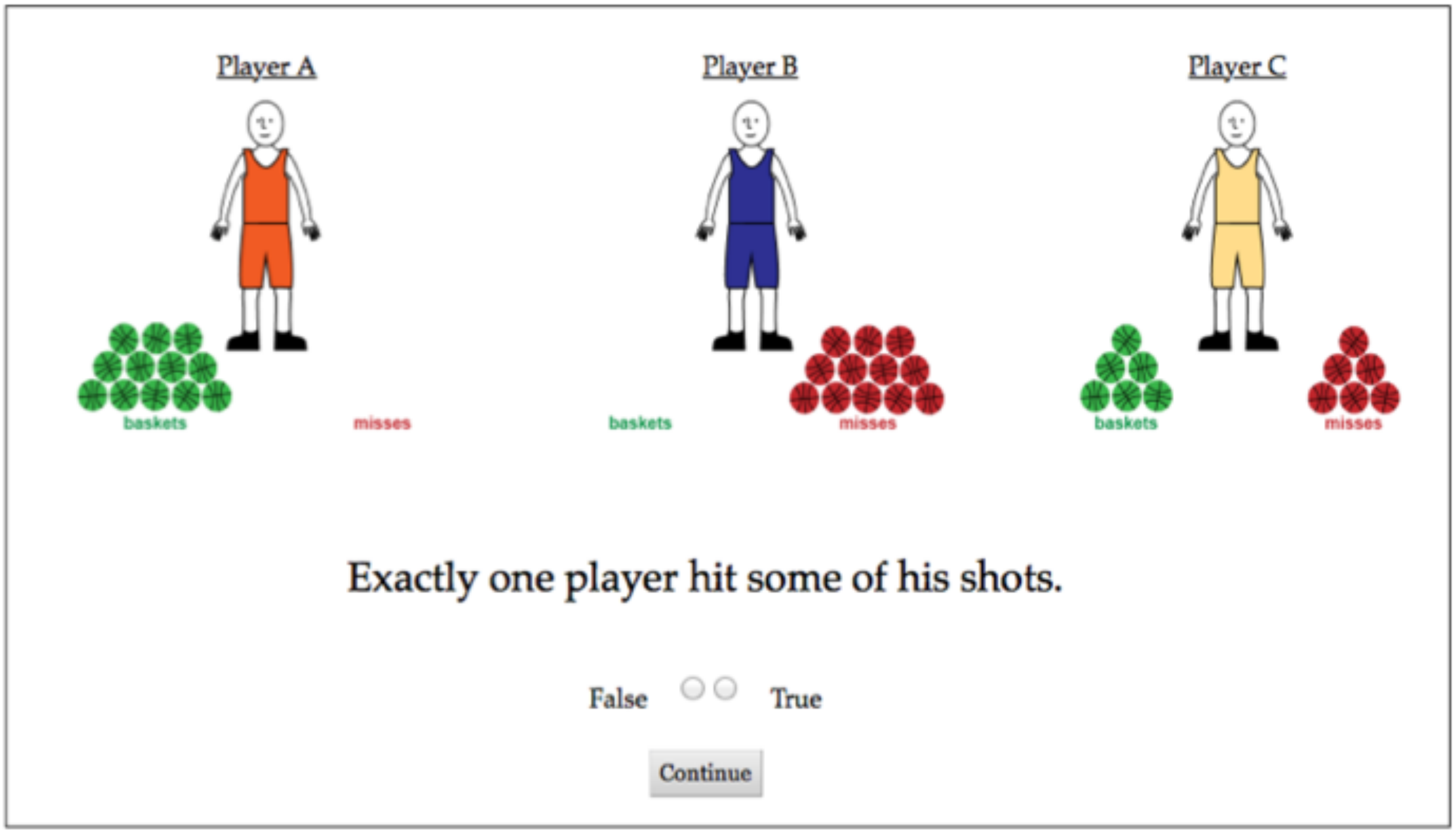
	NN	NS	NA	SN	SS	SA	AN	AS	AA
Player A scored	0	0	0	.24	.19	.16	.18	.16	.07
Player A aced	0	0	0	0	0	0	.36	.30	.34
Player B scored	0	.24	.18	0	.19	.16	0	.16	.07
Player B aced	0	0	.36	0	0	.30	0	0	.34
some player scored	0	.14	.11	.14	.17	.14	.11	.14	.05
some player aced	0	0	.22	0	0	.19	.22	.19	.18
every player scored	0	0	0	0	.31	.27	0	.27	.14
every player aced	0	0	0	0	0	0	0	0	1
no player scored	.31	.14	.12	.14	.06	.05	.12	.05	.01
no player aced	.18	.19	.08	.19	.14	.06	.08	.06	0
0	.01	.01	.32	.01	.01	.15	.32	.15	0

Table 3: Enrichment in the largest space of refinements supported by this lexicon.

	NN	NS	NA	SN	SS	SA	AN	AS	AA
Player A scored	0	0	0	.45	.11	.22	.15	.05	.02
Player A aced	0	0	0	0	0	0	.42	.36	.22
Player B scored	0	.45	.15	0	.11	.05	0	.22	.02
Player B aced	0	0	.42	0	0	.36	0	0	.22
some player scored	0	.25	.09	.25	.06	.12	.09	.12	.01
some player aced	0	0	.24	0	0	.21	.24	.21	.11
every player scored	0	0	0	0	.61	.16	0	.16	.07
every player aced	0	0	0	0	0	0	0	0	1
no player scored	.61	0	.16	0	0	0	.16	0	.06
no player aced	.19	.17	.10	.17	.13	.07	.10	.07	0
0	.15	.13	.13	.13	.10	.09	.13	.09	.05

Table 4: Enrichment using the lexically-driven (neo-Gricean) refinement sets in (14).

Potts et al (2015)



Potts et al (2015)

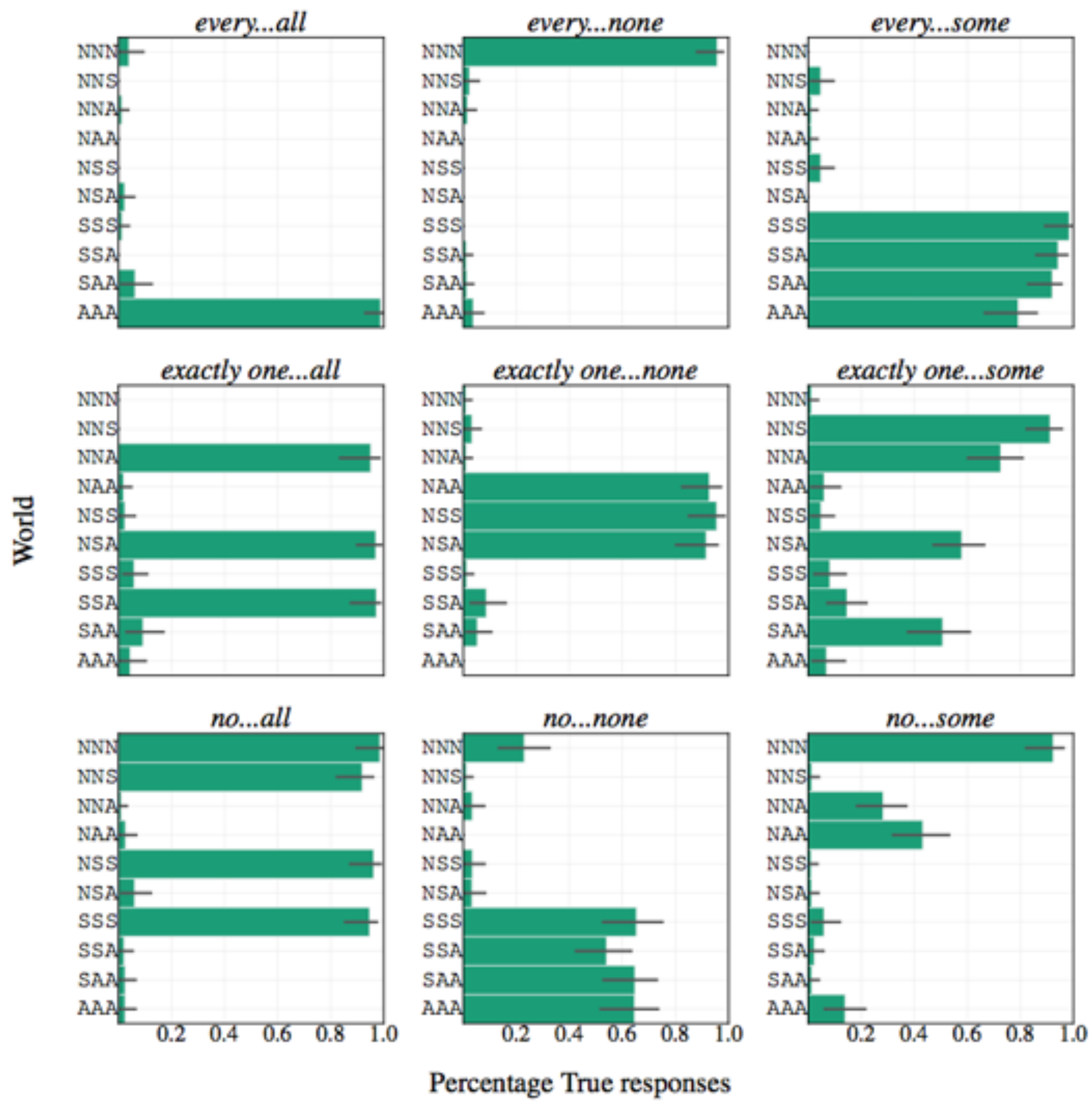
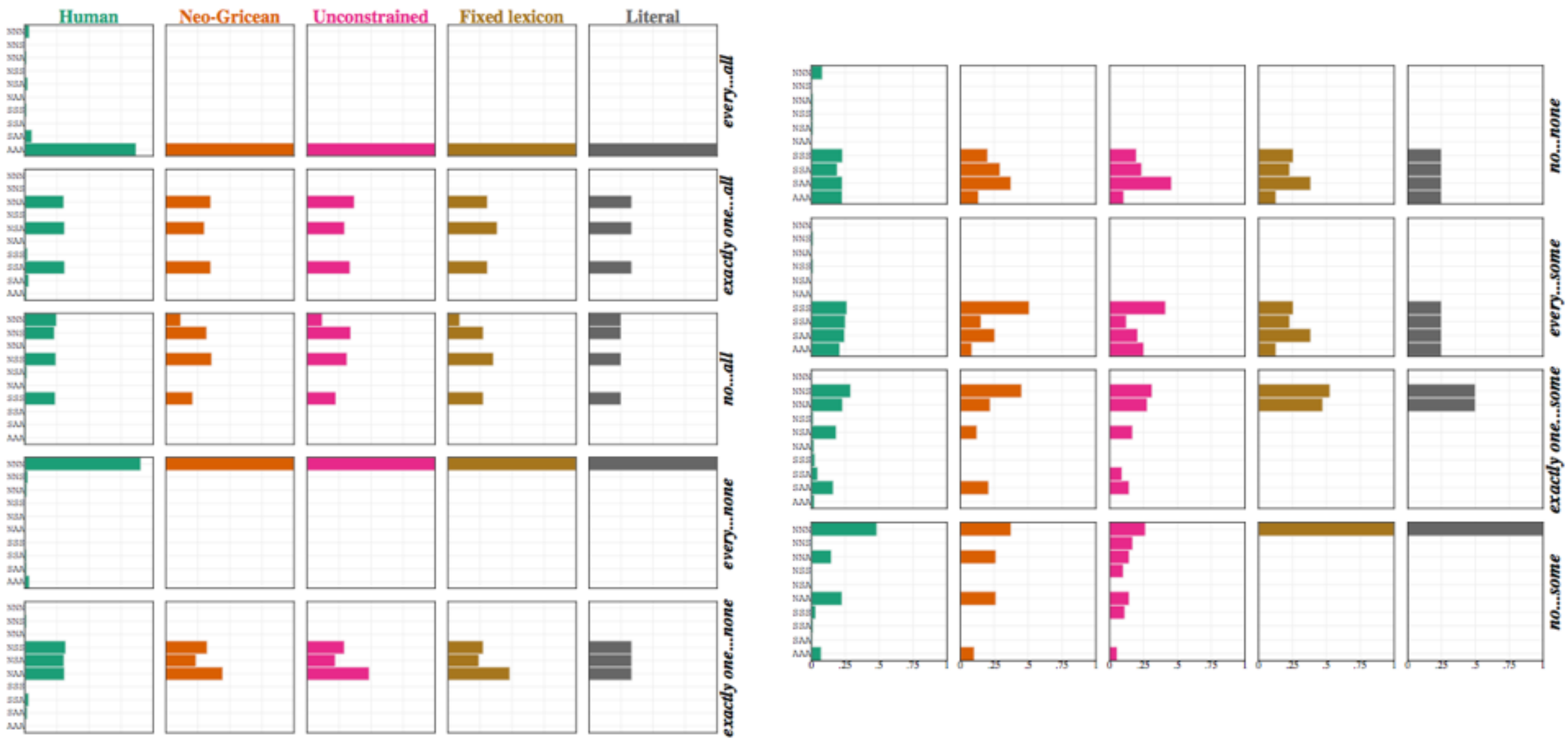


Figure 4: Mean truth-value judgment responses by sentence with bootstrapped 95% confidence intervals.

Potts et al (2015)

	Pearson	Spearman	MSE
Literal semantics	.94	.76	.0065
Fixed-lexicon pragmatics	.92	.76	.0079
Unconstrained uncertainty	.95	.79	.0038
Neo-Gricean uncertainty	.96	.81	.0034

Potts et al (2015)



Exceptional DE contexts

Context: A and B are visiting a resort. B has very particular preferences about the temperature of the springs at the resort: he will bathe in them if they are between 85-95 °F (30-35 °C), or between 105-115 °F (40-45 °C), as he finds the lower temperatures relaxing and the higher temperatures invigorating. A knows about B's preferences, and has checked the water temperature for him.

A: The water isn't warm or scalding.

Exceptional DE contexts

Context: A and B are scientists who study cancer in mice. They are discussing a tumor that one mouse has developed. If it is above 1 mm in size, then it cannot be removed safely. If it is between 0.1-1 mm, then it can be surgically removed, and the mouse can be saved; if it is between 0.01-0.1 mm, then it is too small to be surgically removed, but may still be harmful to the animal; and if it is less than 0.01 mm, then it is so small that it will not harm the animal. These facts about mouse tumors are common knowledge among A and B, and A has gotten some information about the tumor size.

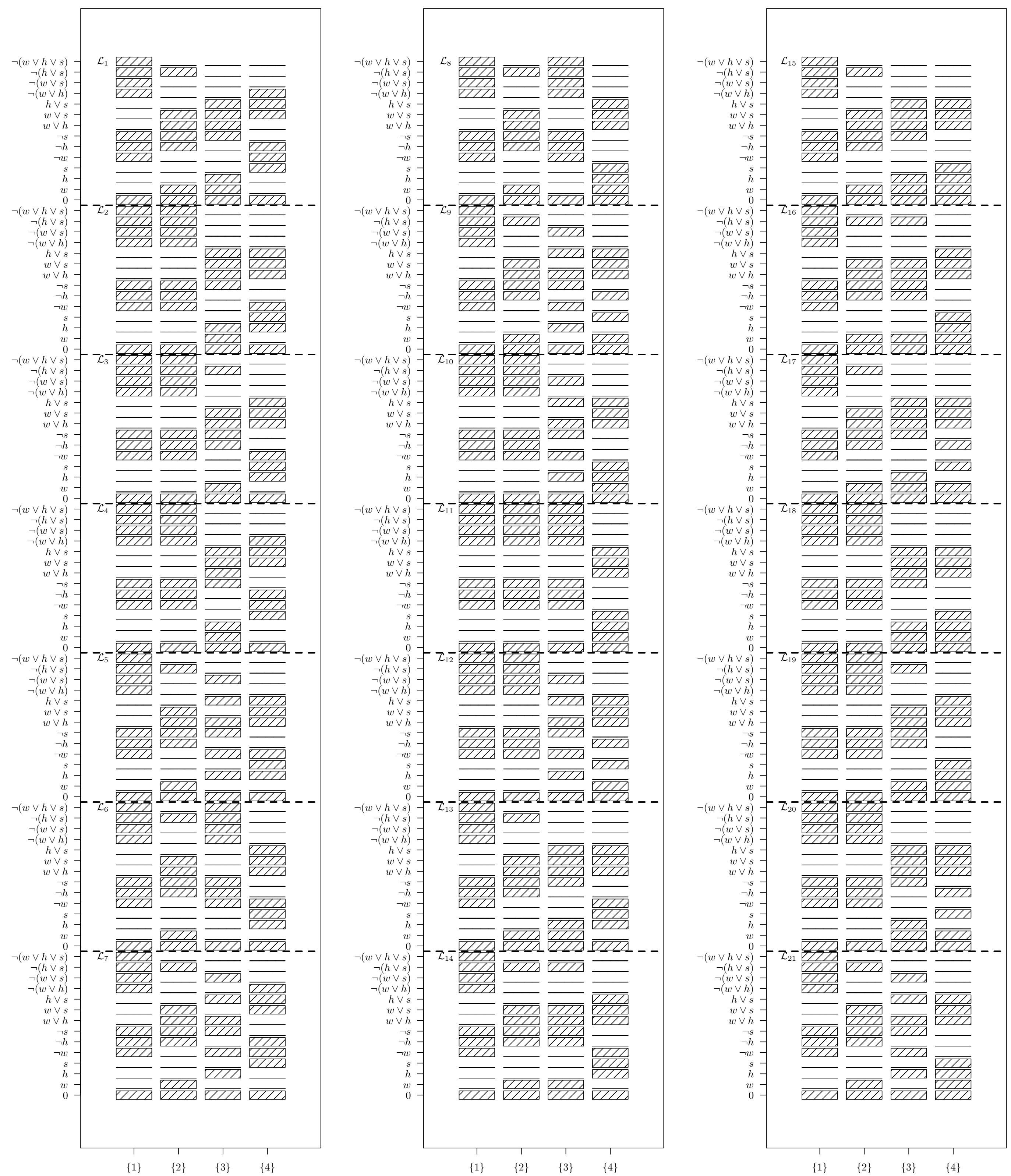
A: The tumor isn't small or microscopic.

Exceptional DE contexts

Strategy for generating the DE implicature:

- First show that “not warm or scalding” will be used when the speaker knows that the water is either cool or hot (but not scalding).
- Then show that “not warm or scalding” will not be used in any other knowledge state.

Exceptional DE contexts



Exceptional DE contexts

Show that “not warm or scalding” will be used when the speaker knows that the water is either cool or hot (but not scalding).

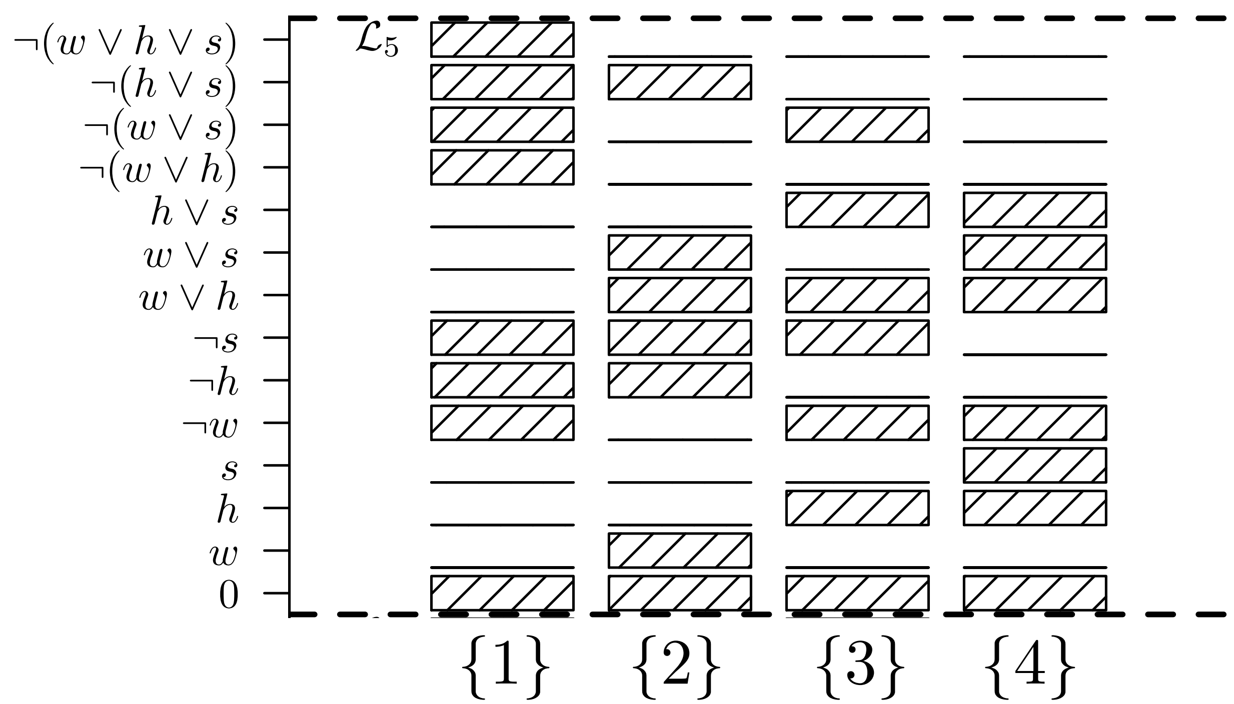
- “Not warm or scalding” is compatible with the knowledge state in 9 lexica.
- These are exactly the lexica in which “Not warm” is compatible with the knowledge state.

Exceptional DE contexts

Show that "not warm or scalding" will be used when the speaker knows that the water is either cool or hot (but not scalding).

- "Not scalding" is always compatible with the knowledge state, but also is compatible with many other knowledge states.
- "Not hot" is compatible in 7 lexica, in which case it means *not scalding*.

Exceptional DE contexts



Exceptional DE contexts

We have shown the first point:

- “Not warm or scalding” will be used when the speaker knows that the water is either cool or hot (but not scalding).
- Then show that “not warm or scalding” will not be used in any other knowledge state.

Exceptional DE contexts

Now we show the second point:

- “Not warm or scalding” will not be used in any other knowledge state.

Exceptional DE contexts

Two crucial knowledge states to consider:

1. The speaker knows that the water is cool.
   - In this case, “not warm or hot” is strictly better than “not warm or scalding.”
2. The speaker knows that the water is hot but not scalding.
   - “Not warm or scalding” is always compatible with the water being cool.
   - “Hot” is usually compatible with the water being hot but not scalding, and often picks out this knowledge state uniquely.

Exceptional DE contexts