Semantic Theory
week 12 – Distributive Situation-state Spaces

Noortje Venhuizen
(based on slides by Harm Brouwer)

Universität des Saarlandes

Summer 2017

Semantics: a psycholinguistic perspective

"charlie plays soccer"

play(charlie,soccer)

Zwaan & Radvansky (1998). *Psychol. Bull.*

Distributed Situation Space (DSS)

- A non-symbolic, distributed representational scheme for meaning
- Situations are represented as vectors in a high-dimensional space called "*situation-state space*"
- DSS vectors capture dependencies between situations, allowing for 'world knowledge'-driven *direct inference*
- To encode all world knowledge, DSS vectors are derived from observations of *states-of-affairs* (situations) in a *microworld*

Frank et al. (2003). *Cogn. Sci.;* Frank et al. (2009). *Cognition*

DSS — The main idea

- Take a snapshot of the world ("a sample") at many different (independent) points in time, and for each snapshot write down the full ***state-of-affairs*** in the world
- Meaning of individual propositions is determined by ***collocation*** with other propositions in full set of states-of-affairs (cf. Distributional Semantics)

Problem: How to record full state-of-affairs in the world for each snapshot?

▸ Limit the scope of the world by using a confined microworld

Defining a Microworld

An ***observation*** (state-of-affairs) in a microworld is defined in terms of the set of ***atomic events***; i.e., each atomic event is either the case or not the case

Class	Variable	Class members (concepts)	#
People	p	charlie, heidi, sophia	3
Games	g	chess, hide&seek, soccer	3
Toys	t	puzzle, ball, doll	3
Places	x	bathroom, bedroom, playground, street	4
Manners of playing	m_{play}	well, badly	2
Manners of winning	m_{win}	easily, difficultly	2
Predicates	—	play, win, lose, place, manner	5

Event name		#
play(p, g)	$3 \times 3 =$	9
play(p, t)	$3 \times 3 =$	9
win(p)		3
lose(p)		3
place(p, x)	$3 \times 4 =$	12
manner(play(p), m_{play})	$3 \times 2 =$	6
manner(win, m_{win})		2
	Total	44

> 2^44 (≈10^13) possible observations, but ***world knowledge*** precludes many

Frank et al. (2009). *Cognition*

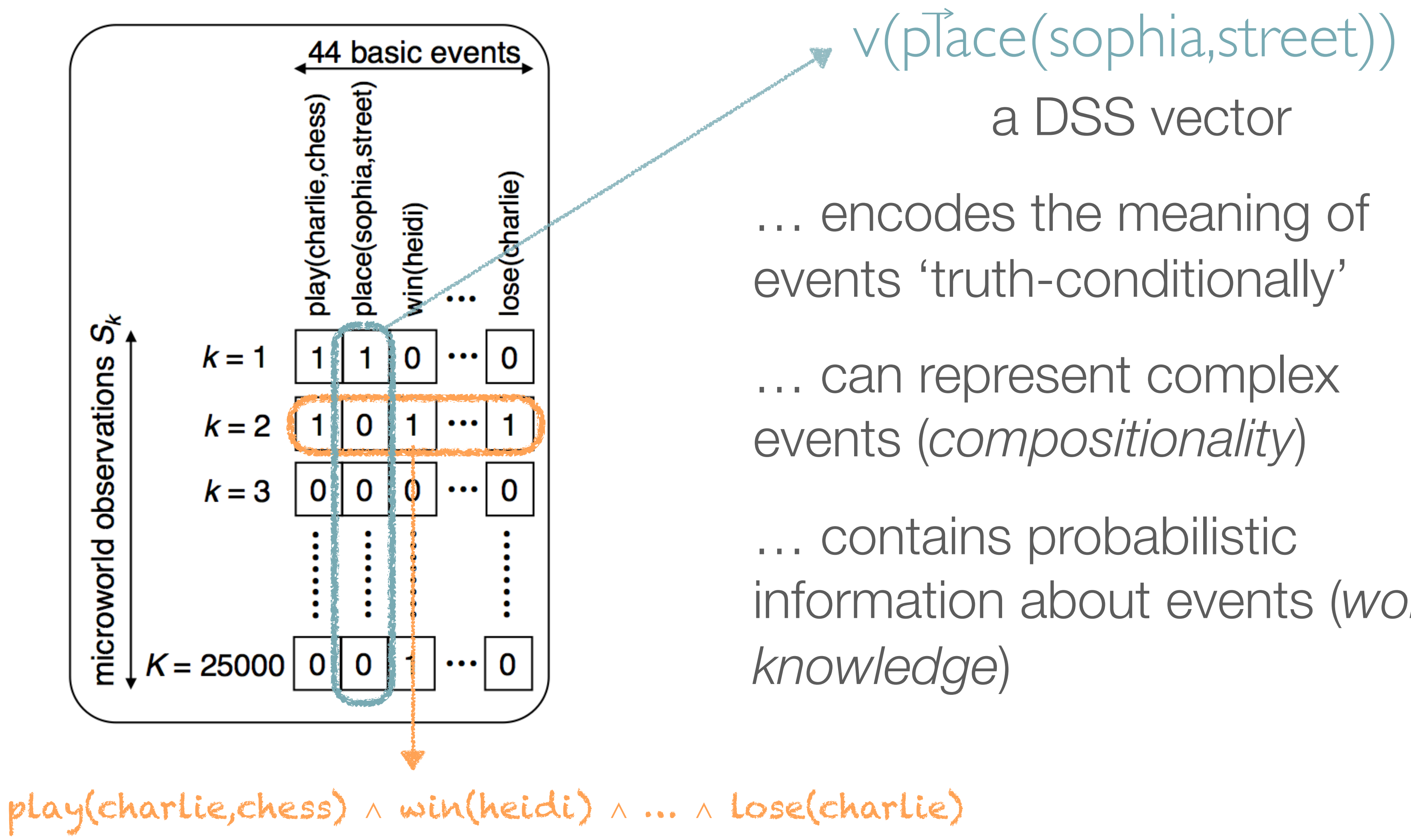
Microworld knowledge

World knowledge enforces constraints on event co-occurrence. Some examples:

- Personal characteristics—each person has a specialty, a preferred toy, and some persons frequent specific places
- Games and toys—each game/toy can only be played (with) in specific places, and has a number of possible player configurations; soccer is played with a ball
- Being there—everybody is exactly at one place; if hide&seek is played in the playground, all players are there; all chess players are in the same place
- Winning and losing—only one can win, and one cannot win and lose; if someone wins, all other players lose

Frank et al. (2009). *Cognition*

Distributed Situation-state Space



a DSS vector

… encodes the meaning of events 'truth-conditionally'

… can represent complex events (*compositionality*)

… contains probabilistic information about events (*world knowledge*)

DSS: Compositionality

The DSS vectors of ***atomic events*** are the columns of the DSS matrix

The DSS vectors of ***complex events*** can be found through (fuzzy) propositional logic:

$$\vec{v}(\neg a) = 1 - \vec{v}(a)$$

$$\vec{v}(a \wedge b) = \vec{v}(a)\vec{v}(b) \quad \text{where} \quad \vec{v}(a \wedge a) = \vec{v}(a)$$

Which gives us $\vec{v}(a \uparrow b) = \vec{v}(\neg\vec{v}(a \wedge b))$ and hence *functional completeness*:

$$\vec{v}(a \vee b) = \vec{v}(\vec{v}(a \uparrow a) \uparrow \vec{v}(b \uparrow b))$$

$$\vec{v}(a \rightarrow b) = \vec{v}(a \uparrow \vec{v}(b \uparrow b)) = \vec{v}(a \uparrow \vec{v}(a \uparrow b))$$

$$\vec{v}(a \veebar b) = \vec{v}(\vec{v}(a \uparrow \vec{v}(a \uparrow b)) \uparrow \vec{v}(b \uparrow \vec{v}(a \uparrow b)))$$

> allows for deriving DSS vectors for events of arbitrary logical complexity

DSS: Probabilistic information

Situation vectors encode events by means of *co-occurrence probabilities*

Prior belief in atomic event *a* (= estimate of its probability):

$$B(a) = \frac{1}{k}\sum_i \vec{v}_i(a) \approx Pr(a)$$

Prior conjunction belief of atomic events *a* and *b*:

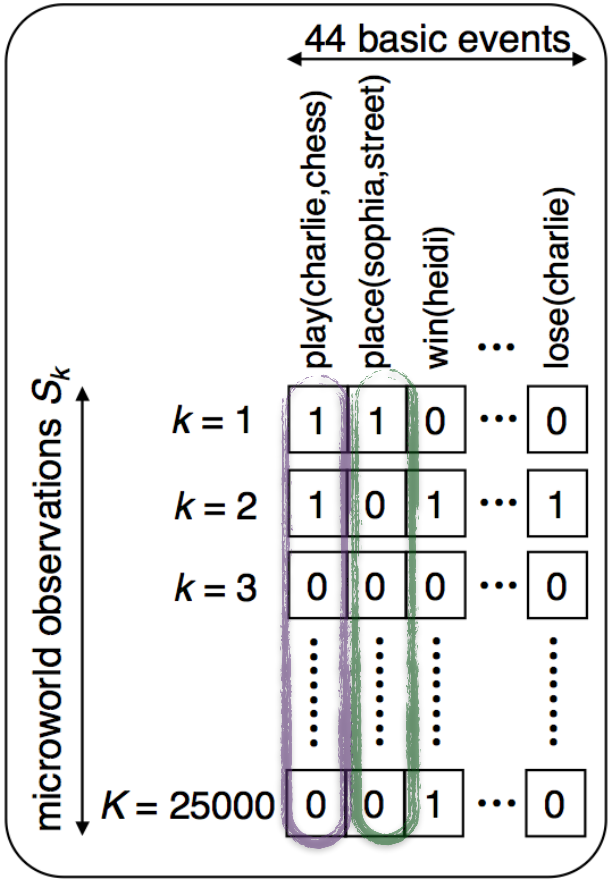
$$B(a \wedge b) = \frac{1}{k}\sum_i \vec{v}_i(a)\vec{v}_i(b) \approx Pr(a \wedge b) \quad \text{where} \quad B(a \wedge a) = B(a)$$

Prior conditional belief of atomic event *a* given *b*:

$$B(a|b) = \frac{B(a \wedge b)}{B(b)} \approx Pr(a|b)$$

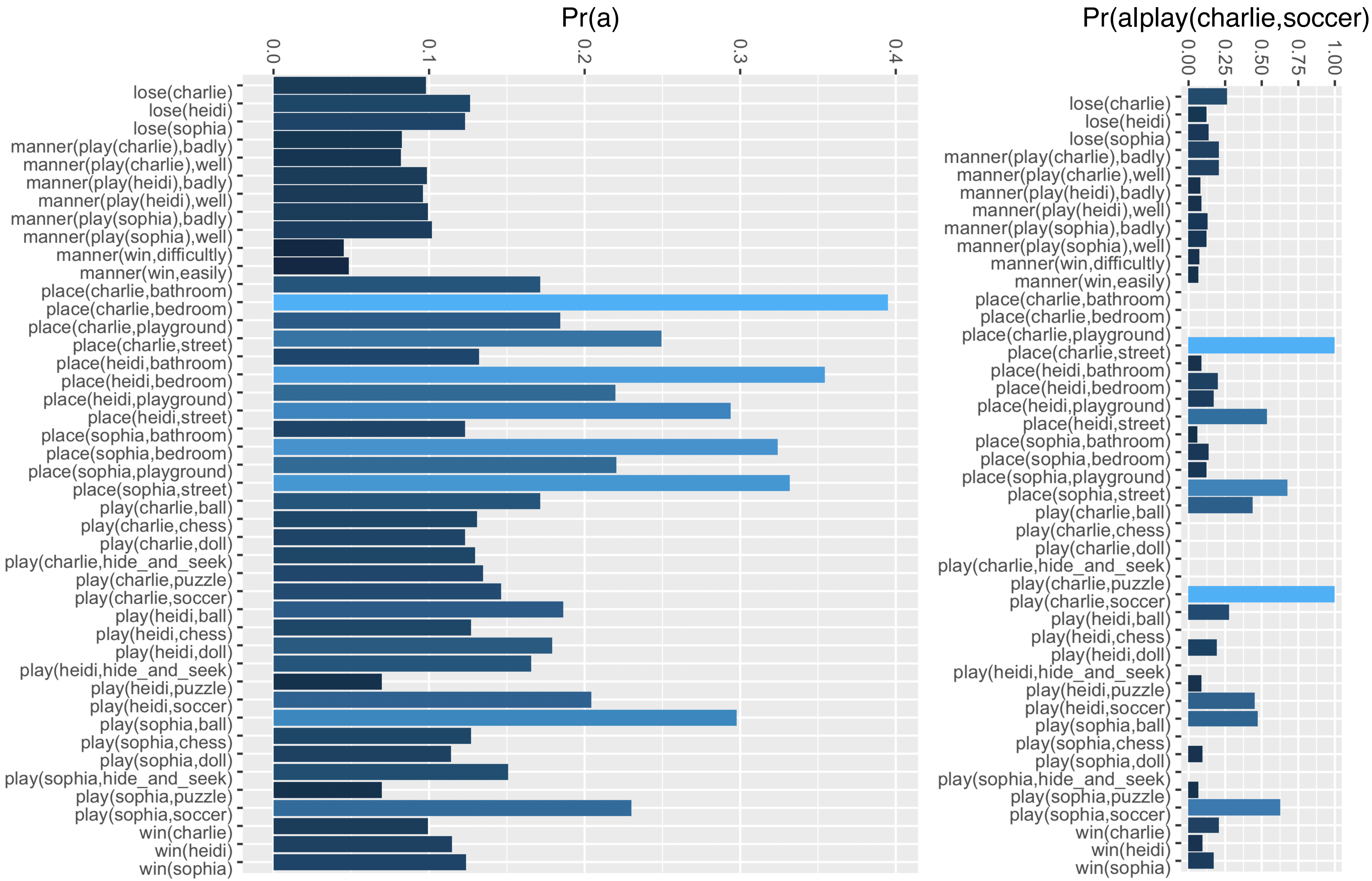
Critically, *a* and *b* can be atomic or complex events

$B(a|b) \approx Pr(a|b)$ means $\vec{v}(b)$ encodes *b* and all that depends upon *b*; this allows 'world knowledge'-driven inference



Frank et al. (2009). *Cognition*

Microworld Probabilities



Quantifying comprehension

Beyond conditional belief—how much is event *a* ‘understood’ from event *b*

- Knowing *b* ***increases*** belief in *a*: the conditional belief *B(a|b)* is higher than the prior belief *B(a)*

- Knowing *b* ***decreases*** belief in *a*: the conditional belief *B(a|b)* is lower than the prior belief *B(a)*

$$comprehension(a,b) = \begin{cases} \frac{B(a|b)-B(a)}{1-B(a)} & \text{if } B(a|b) > B(a) \\ \frac{B(a|b)-B(a)}{B(a)} & \text{otherwise} \end{cases}$$

-1 ≤ *comprehension(a,b)* ≤ +1

+1 = perfect positive comprehension: b took away all uncertainty in a

-1 = perfect negative comprehension: b took away all certainty in a

Frank et al. (2009). *Cognition*


Comprehension scores
Atomic event b
Atomic event a
value
1.0
0.5
0.0
−0.5
−1.0
win(sophia)
win(heidi)
win(charlie)
play(sophia,soccer)
play(sophia,puzzle)
play(sophia,hide_and_seek)
play(sophia,doll)
play(sophia,chess)
play(sophia,ball)
play(heidi,soccer)
play(heidi,puzzle)
play(heidi,hide_and_seek)
play(heidi,doll)
play(heidi,chess)
play(heidi,ball)
play(charlie,soccer)
play(charlie,puzzle)
play(charlie,hide_and_seek)
play(charlie,doll)
play(charlie,chess)
play(charlie,ball)
place(sophia,street)
place(sophia,playground)
place(sophia,bedroom)
place(sophia,bathroom)
place(heidi,street)
place(heidi,playground)
place(heidi,bedroom)
place(heidi,bathroom)
place(charlie,street)
place(charlie,playground)
place(charlie,bedroom)
place(charlie,bathroom)
manner(win,easily)
manner(win,difficulty)
manner(play(sophia),well)
manner(play(sophia),badly)
manner(play(heidi),well)
manner(play(heidi),badly)
manner(play(charlie),well)
manner(play(charlie),badly)
lose(sophia)
lose(heidi)
lose(charlie)

From sentences to vectors

Use FOL as an intermediate representation for sentences; apply composition rules on DSS vectors to arrive at complex DSS vector

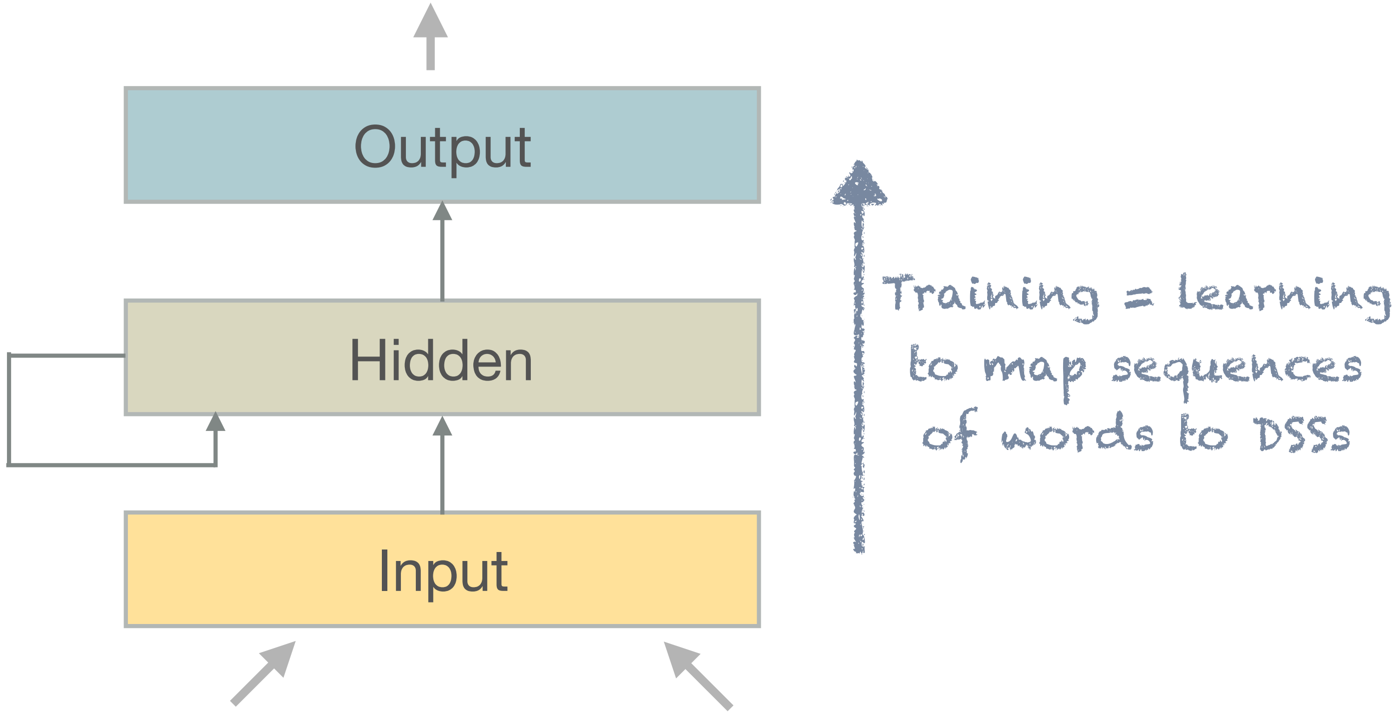
charlie plays chess	play(c, chess)
chess is played by charlie	play(c, chess)
girl plays chess	play(h, chess) ∨ play(s, chess)
heidi plays game	play(h, chess) ∨ play(h, hide&seek) ∨ play(h, soccer)
heidi plays with toy	play(h, puzzle) ∨ play(h, ball) ∨ play(h, doll)
sophia plays soccer well	play(s, soccer) ∧ manner(play(s), well)
sophia plays with ball in street	play(s, ball) ∧ place(s, street)
someone plays with doll	play(c, doll) ∨ play(h, doll) ∨ play(s, doll)
doll is played with	play(c, doll) ∨ play(h, doll) ∨ play(s, doll)
charlie plays	play(c, chess) ∨ play(c, hide&seek) ∨ play(c, soccer) ∨ play(c, puzzle) ∨ play(c, ball) ∨ play(c, doll)

Frank et al. (2009). *Cognition*

Applying DSS in a neural network model

play(charlie,chess) ∧ place(charlie,bathroom)

[1,1,1,0,0,1,1,0,1,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,0,0,0,1,1,1,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,
0,0,0,0,0,1,1,1,0,1,0,0,0,1,1,0,0,1,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,1,0,1,0,1,0,0,0,0,1,0,1,0,0,0,0,0,…]



[0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0], [0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0], …, …

“charlie”, “plays”, “chess”, “in”, “the”, “bathroom”

charlie plays chess in the bathroom

Word-by-word inferencing

model3 — rlwrap — 117×55

```
model:all_sents> dssScores basic_events "charlie plays chess"

**** Sentence:  charlie plays chess
**** Semantics: play(charlie,chess)
****
****                            charlie            plays              chess
****
****                            +0.08693  -0.01071  +0.07622  +0.72407  +0.80029
****
**** play(charlie,chess)        +0.08693  -0.01071  +0.07622  +0.72407  +0.80029  play(charlie,chess)
**** play(charlie,hide_and_seek) +0.04279  -0.01568  +0.02710  -0.77386  -0.74676  play(charlie,hide_and_seek)
**** play(charlie,soccer)       +0.12169  -0.03329  +0.08841  -0.87402  -0.78562  play(charlie,soccer)
**** play(heidi,chess)          +0.01111  +0.00486  +0.01597  +0.41746  +0.43343  play(heidi,chess)
**** play(heidi,hide_and_seek)  -0.08301  -0.00589  -0.08890  -0.77328  -0.86218  play(heidi,hide_and_seek)
**** play(heidi,soccer)         +0.00730  -0.00709  +0.00021  -0.88450  -0.88429  play(heidi,soccer)
**** play(sophia,chess)         +0.00809  +0.00035  +0.00844  +0.35767  +0.36612  play(sophia,chess)
**** play(sophia,hide_and_seek) -0.11984  -0.00710  -0.12694  -0.71115  -0.83810  play(sophia,hide_and_seek)
**** play(sophia,soccer)        +0.03251  -0.02039  +0.01211  -0.90027  -0.88816  play(sophia,soccer)
**** play(charlie,puzzle)       -0.03140  +0.04140  +0.01001  -0.69627  -0.68626  play(charlie,puzzle)
**** play(charlie,ball)         +0.01103  -0.00114  +0.00989  -0.81493  -0.80505  play(charlie,ball)
**** play(charlie,doll)         -0.23204  +0.11227  -0.11977  -0.67018  -0.78995  play(charlie,doll)
**** play(heidi,puzzle)         -0.12599  +0.04865  -0.07733  -0.12012  -0.19746  play(heidi,puzzle)
**** play(heidi,ball)           +0.00373  -0.00318  +0.00055  -0.37581  -0.37527  play(heidi,ball)
**** play(heidi,doll)           +0.00341  -0.00959  -0.00618  -0.20651  -0.21269  play(heidi,doll)
**** play(sophia,puzzle)        -0.04606  -0.03433  -0.08039  +0.08060  +0.00021  play(sophia,puzzle)
**** play(sophia,ball)          +0.00915  -0.00004  +0.00912  -0.40594  -0.39683  play(sophia,ball)
**** play(sophia,doll)          -0.03115  +0.02611  -0.00504  -0.07983  -0.08487  play(sophia,doll)
**** win(charlie)               +0.09208  -0.03814  +0.05393  +0.19446  +0.24839  win(charlie)
**** win(heidi)                 +0.01076  -0.10392  -0.09316  -0.11974  -0.21290  win(heidi)
**** win(sophia)                +0.01201  -0.06818  -0.05617  -0.34537  -0.40154  win(sophia)
**** lose(charlie)              +0.08350  -0.04772  +0.03578  -0.00125  +0.03453  lose(charlie)
**** lose(heidi)                +0.01591  -0.01308  +0.00283  +0.05884  +0.06167  lose(heidi)
**** lose(sophia)               +0.02748  -0.02746  +0.00001  +0.06213  +0.06214  lose(sophia)
**** place(charlie,bathroom)    -0.45570  -0.04047  -0.49617  -0.40434  -0.90052  place(charlie,bathroom)
**** place(charlie,bedroom)     +0.11246  +0.01725  +0.12972  +0.71078  +0.84050  place(charlie,bedroom)
**** place(charlie,playground)  -0.24599  +0.07566  -0.17033  -0.65517  -0.82550  place(charlie,playground)
**** place(charlie,street)      +0.07425  -0.02316  +0.05109  -0.86119  -0.81011  place(charlie,street)
**** place(heidi,bathroom)      -0.01684  +0.02099  +0.00415  -0.44904  -0.44489  place(heidi,bathroom)
**** place(heidi,bedroom)       -0.00530  +0.00981  +0.00451  +0.45972  +0.46424  place(heidi,bedroom)
**** place(heidi,playground)    -0.02113  -0.01805  -0.03918  -0.29745  -0.33664  place(heidi,playground)
**** place(heidi,street)        +0.01236  -0.00930  +0.00307  -0.57345  -0.57038  place(heidi,street)
**** place(sophia,bathroom)     -0.03025  +0.03582  +0.00558  -0.34721  -0.34164  place(sophia,bathroom)
**** place(sophia,bedroom)      -0.01583  -0.00878  -0.02461  +0.43567  +0.41106  place(sophia,bedroom)
**** place(sophia,playground)   -0.05556  +0.01669  -0.03886  -0.15951  -0.19837  place(sophia,playground)
**** place(sophia,street)       +0.03195  -0.01436  +0.01759  -0.60041  -0.58283  place(sophia,street)
**** manner(play(charlie),well)  +0.04509  -0.00978  +0.03531  +0.05307  +0.08838  manner(play(charlie),well)
**** manner(play(charlie),badly) +0.04845  -0.01236  +0.03610  +0.05926  +0.09535  manner(play(charlie),badly)
**** manner(play(heidi),well)   -0.01155  +0.00267  -0.00888  +0.01330  +0.00442  manner(play(heidi),well)
**** manner(play(heidi),badly)  -0.02207  +0.00191  -0.02016  +0.01801  -0.00215  manner(play(heidi),badly)
**** manner(play(sophia),well)  +0.00287  -0.01083  -0.00797  -0.13684  -0.14481  manner(play(sophia),well)
**** manner(play(sophia),badly) +0.00391  -0.00347  +0.00044  -0.02596  -0.02552  manner(play(sophia),badly)
**** manner(win,easily)         +0.01757  -0.01269  +0.00488  +0.03096  +0.03583  manner(win,easily)
**** manner(win,difficultly)    +0.01483  -0.01074  +0.00409  +0.01167  +0.01576  manner(win,difficultly)

model:all_sents>
```

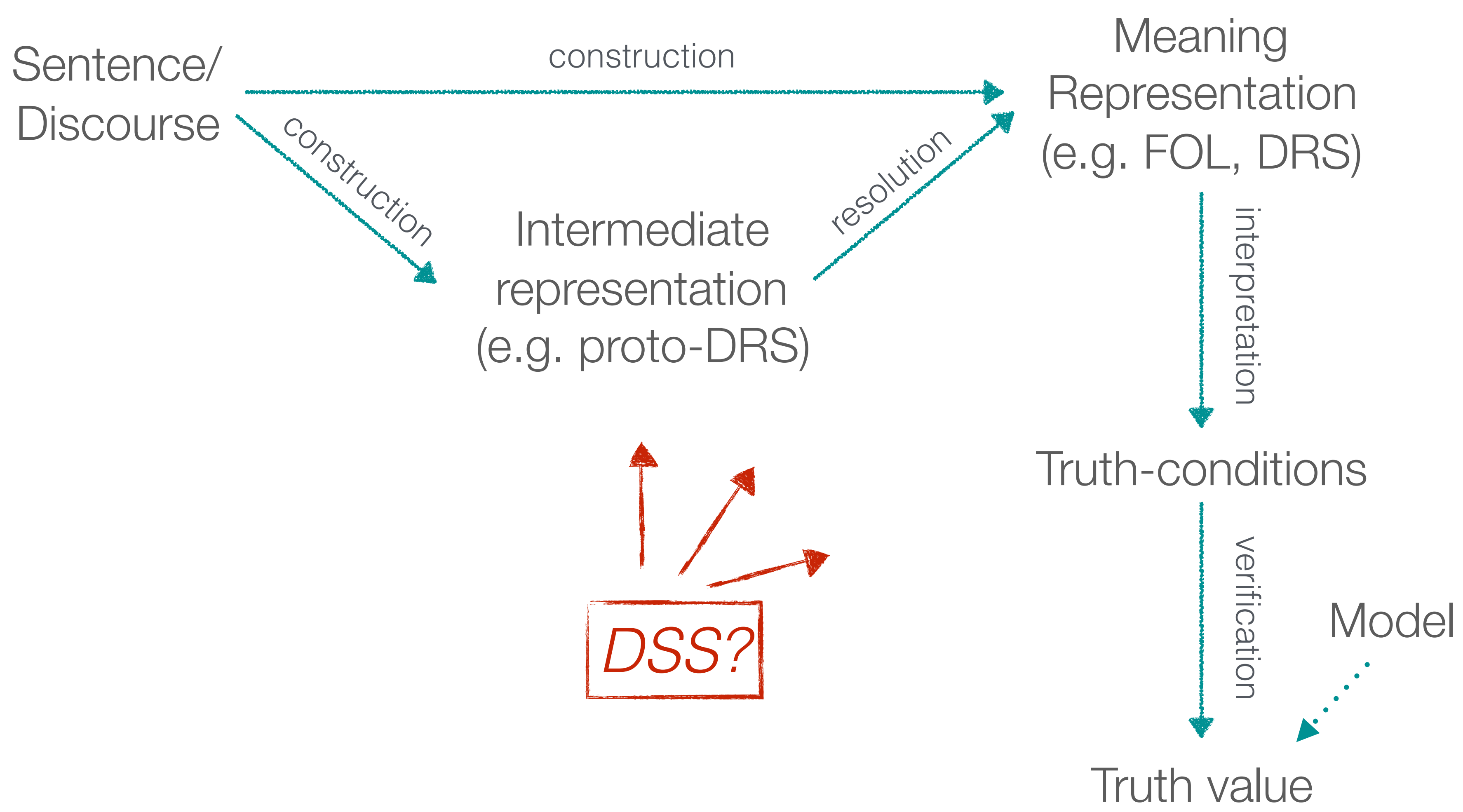
DSS evaluation

The DSS representations are…

- *Neurally plausible*—can be implemented at the neural level (e.g., in a neural network model)
- *Expressive*—capture various aspects of meaning, e.g., negation, quantification
- *Compositional*—meaning of complex propositions is derived from the meaning of their parts
- *Graded*—capture probabilistic dependencies between propositions
- *Inferential*—capture inferences that go beyond literal propositional content
- *Incremental*—can be constructed on a word-by-word basis

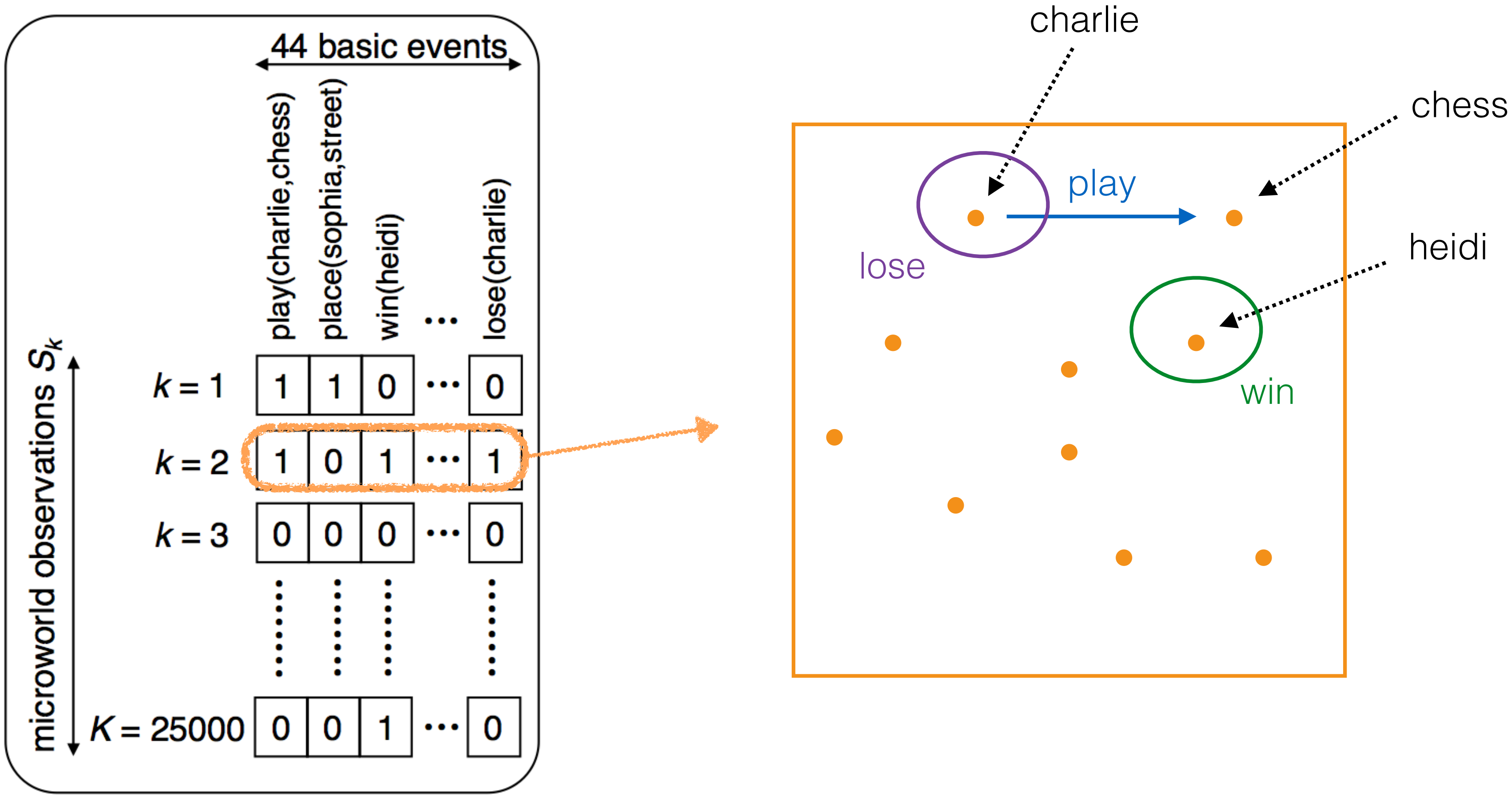
Brouwer et al. (2017). *Festschrift for John Nerbonne*

Back to Semantic Theory



DSSs as collections of logical models

- Each observation in a DSS (i.e., each row in the matrix) represents a *logical model*



DSSs as collections of logical models (cont.)

- Each observation in a DSS (i.e., each row in the matrix) represents a *logical model*
- A set of observations is a collection of models that describes *possible states-of-affairs* in the world (ideally exhaustively, i.e., all *lawful* configurations of atomic events)
- This provides logical models with a *probabilistic dimension*
- DSS observations should in principle be able to capture all *formal properties* that logical models can → How?

Back to: Generalized Quantifiers

Bill $\mapsto \lambda P.P(b^*)$

- $[\![\text{Bill}]\!]^M = \{ P \subseteq U_M \mid b^* \in P\}$ ~ *“the set of properties P, such that Bill is P”*

$[\![\text{charlie}]\!]^{DSS} = \mathrm{U}(\text{event}(\text{charlie})) = [\![\text{play}(\text{charlie},\text{chess})]\!]^{DSS} \vee [\![\text{win}(\text{charlie})]\!]^{DSS} \vee \ldots$

1	1	1
1	0	1
0	0	0
0	0	0
1	1	0
…	…	…

~ *“the set of observations O, such that Charlie does something in O”*

Back to: Presuppositions

(1) *Charlie managed to win at chess*
≫ Charlie tried to win at chess

How to capture this in DSS?

- Add **basic events**: manage(charlie,win) & try(charlie,win)
- Add **world knowledge**: each observation that contains manage(charlie,win) or ¬manage(charlie,win) should also contain try(charlie,win).

Result: $[\![\text{manage(charlie,win)}]\!]^{DSS}$, $[\![\neg\text{manage(charlie,win)}]\!]^{DSS}$, $[\![\text{try(charlie,win)}]\!]^{DSS}$

0	1	1	0	0	...

1	0	0	1	1	...

1	1	1	1	1	...

???

How to fix this?

DSS and Semantic Theory: open questions

How to capture other formal aspects of meaning?

- Lexical inferences
- Quantifier scope
- Monotonicity
- Event structure
- Temporal aspects
- Anaphoric reference
- …