

Simulating Language

Lecture 8: What is the constructor bias?

Simon Kirby

simon@ling.ed.ac.uk



Smith's (2002) approach to language evolution

- Previous approaches:

Build learners with particular biases

Test them with particular hand-built language
(this is the **acquisition** test)

- But where do the languages come from?
- Cultural evolution through iterated learning
Learners learn from other learners in a population
- Two new tests of learning bias in a population:

Maintenance test

Construction test

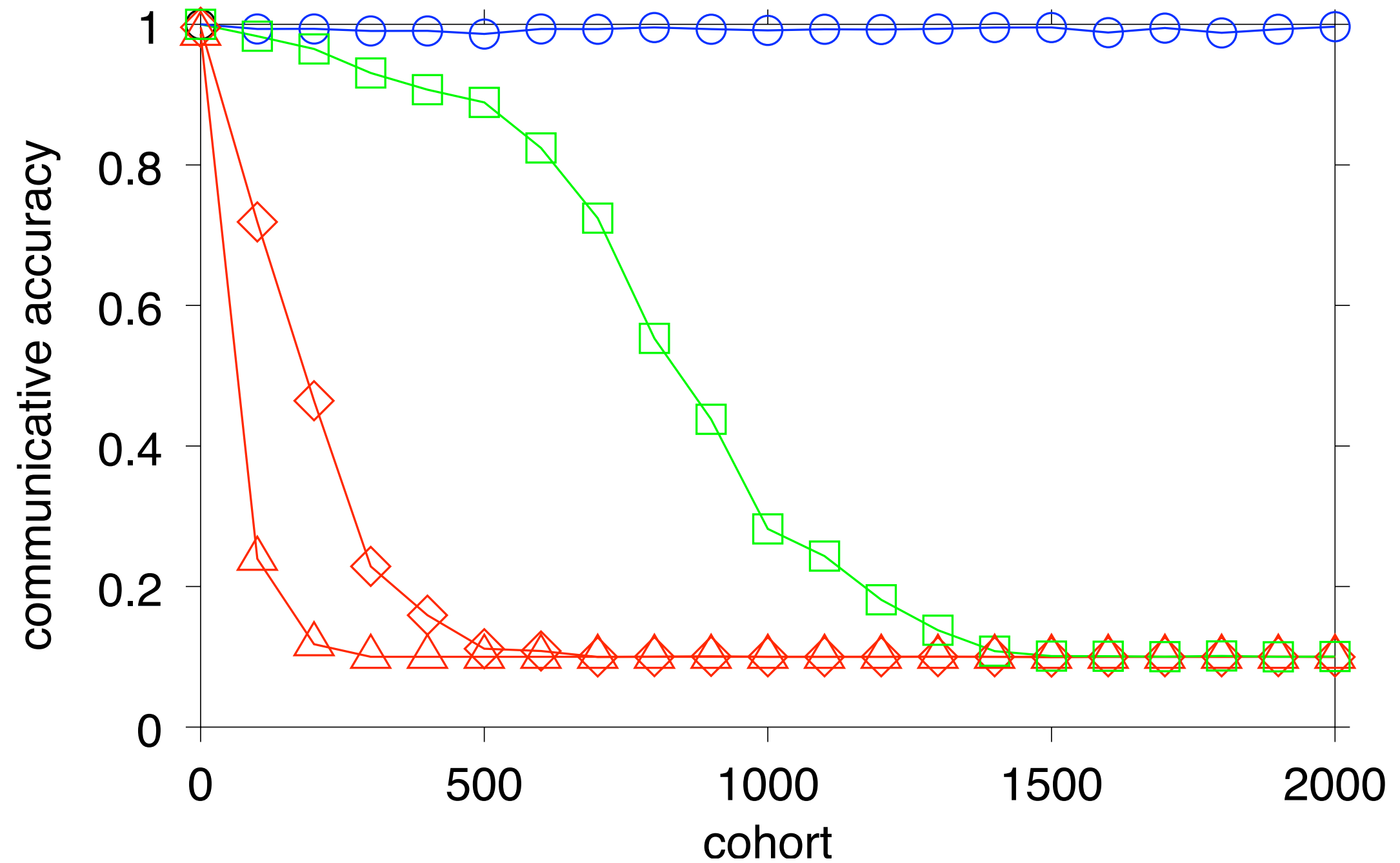
The acquisition test results

- If we look at -1, 0, or 1 for α , β , γ and δ , then there are 81 learning rules
- 50 of these fail the **acquisition** test. We will call these **non-learners**
- 31 pass the test: call these **learners**

For all learners: $\alpha + \delta > \beta + \gamma$

For all non-learners: $\alpha + \delta \leq \beta + \gamma$

Maintenance test results

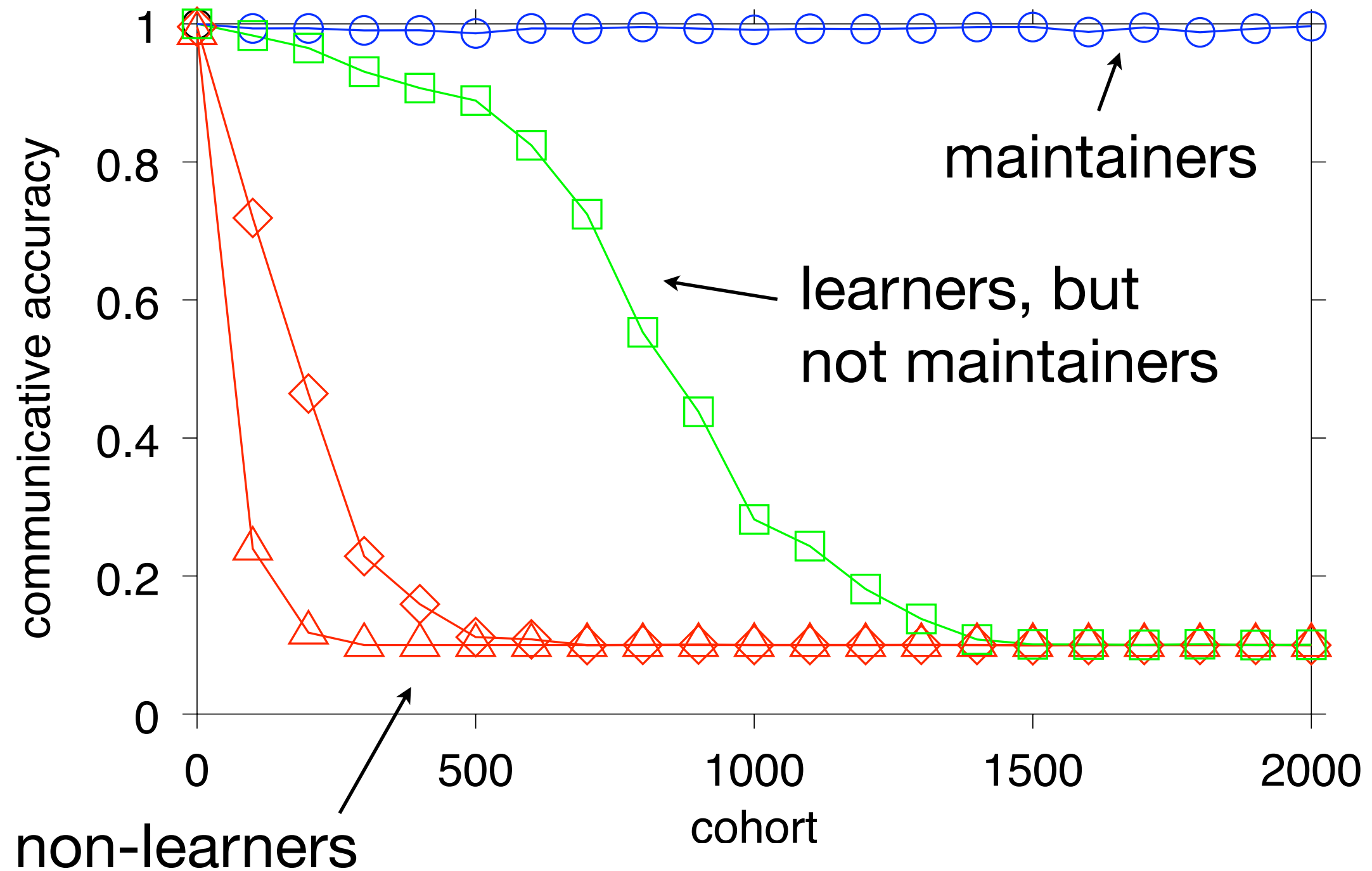


Maintenance test results

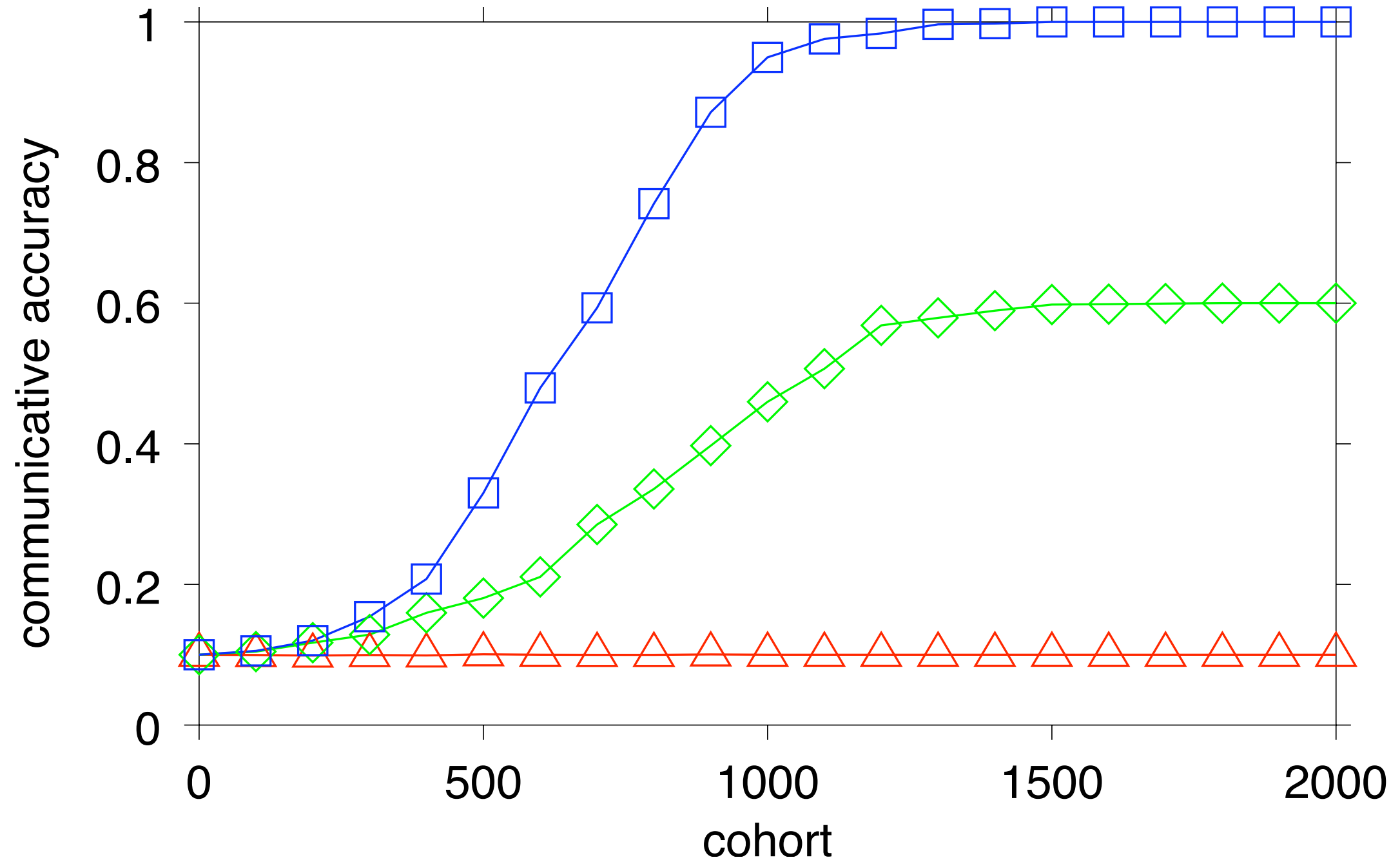
- Out of the 81 rules:
 - 63 fail the **maintenance** test
 - 18 pass: call them **maintainers**
Note, these are a subset of learners

For all maintainers: $\alpha > \beta$ & $\delta \geq \gamma$

Maintenance test results



Construction test results



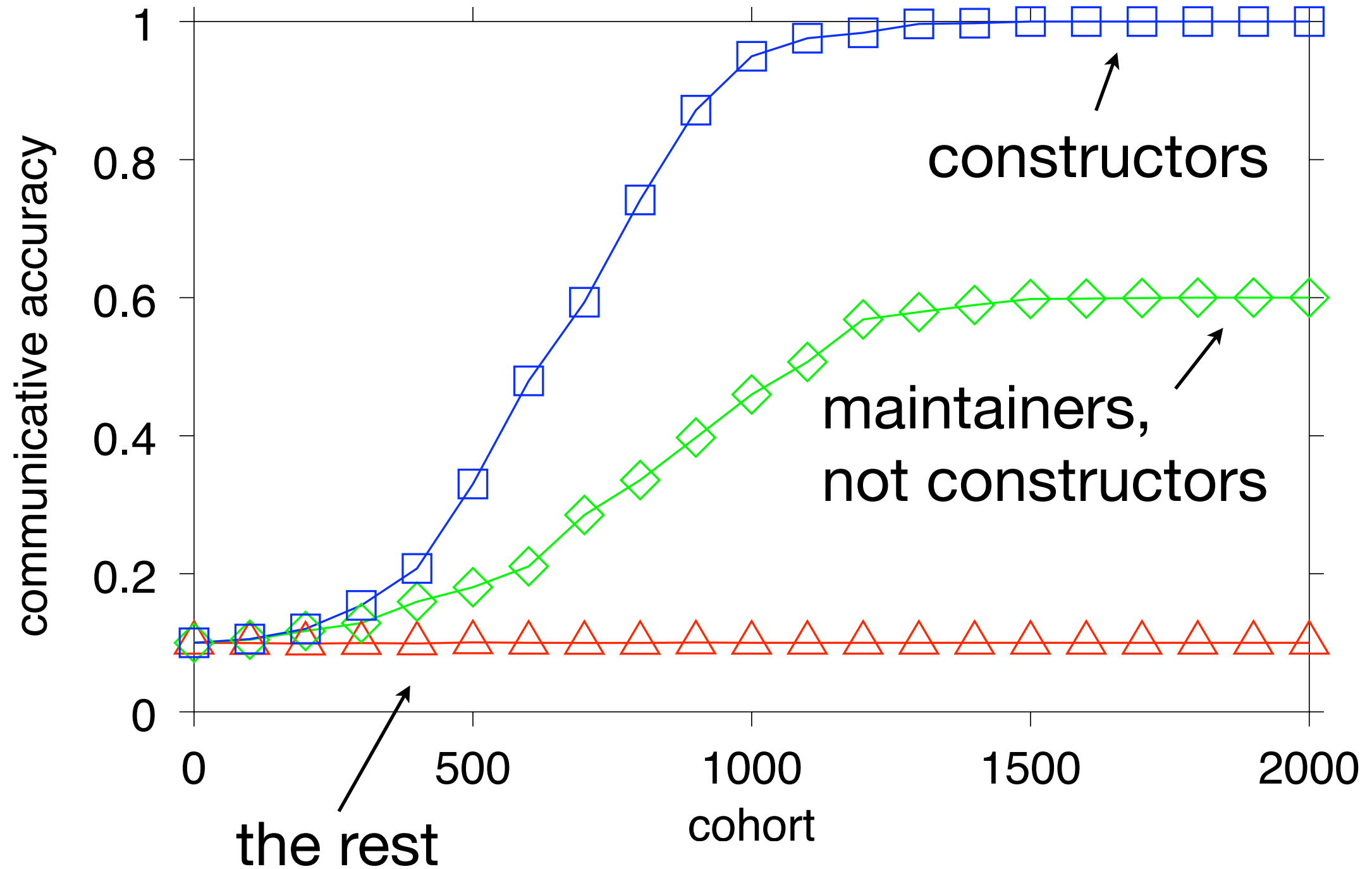
Construction test results

- Out of the 81 rules:
 - 72 fail the construction test
 - 9 pass: call these **constructors**
These are a subset of the maintainers

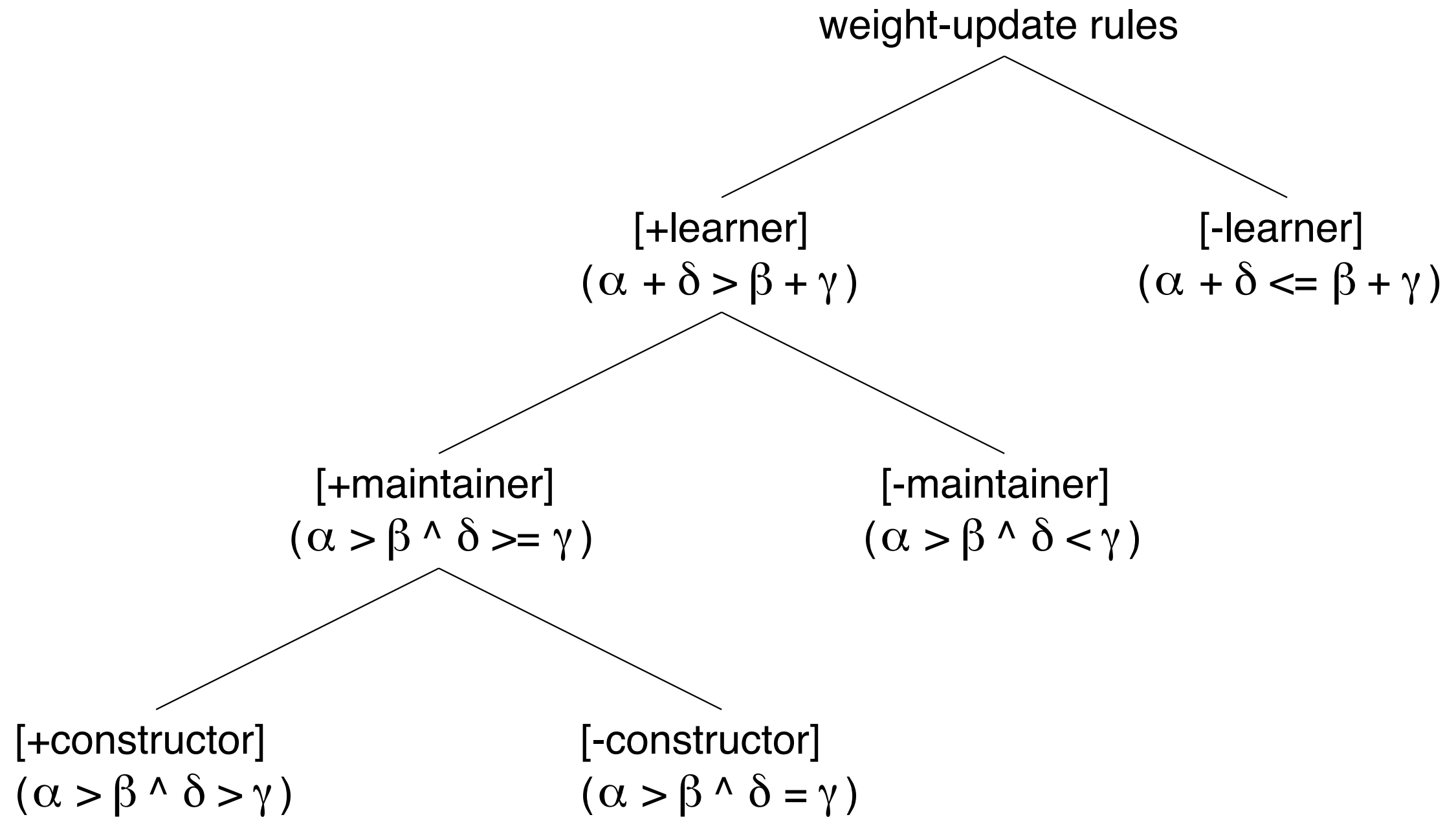
For all constructors: $\alpha > \beta$ & $\delta > \gamma$

(sometimes this is called “lateral inhibition”)

Construction test results



A hierarchy



Bias

- Different weight update rules correspond to different ways of learning
- They come with different *biases*
- Population's language (in this case, just a vocabulary really) evolves to fit these biases
- Biases are a consequence of α , β , γ and δ
- But what exactly are these different biases?
- How do they relate to the *human* vocabulary learning strategy?

Three patterns

learning: $\alpha + \delta > \beta + \gamma$

maintenance: $\alpha > \beta$ & $\delta \geq \gamma$

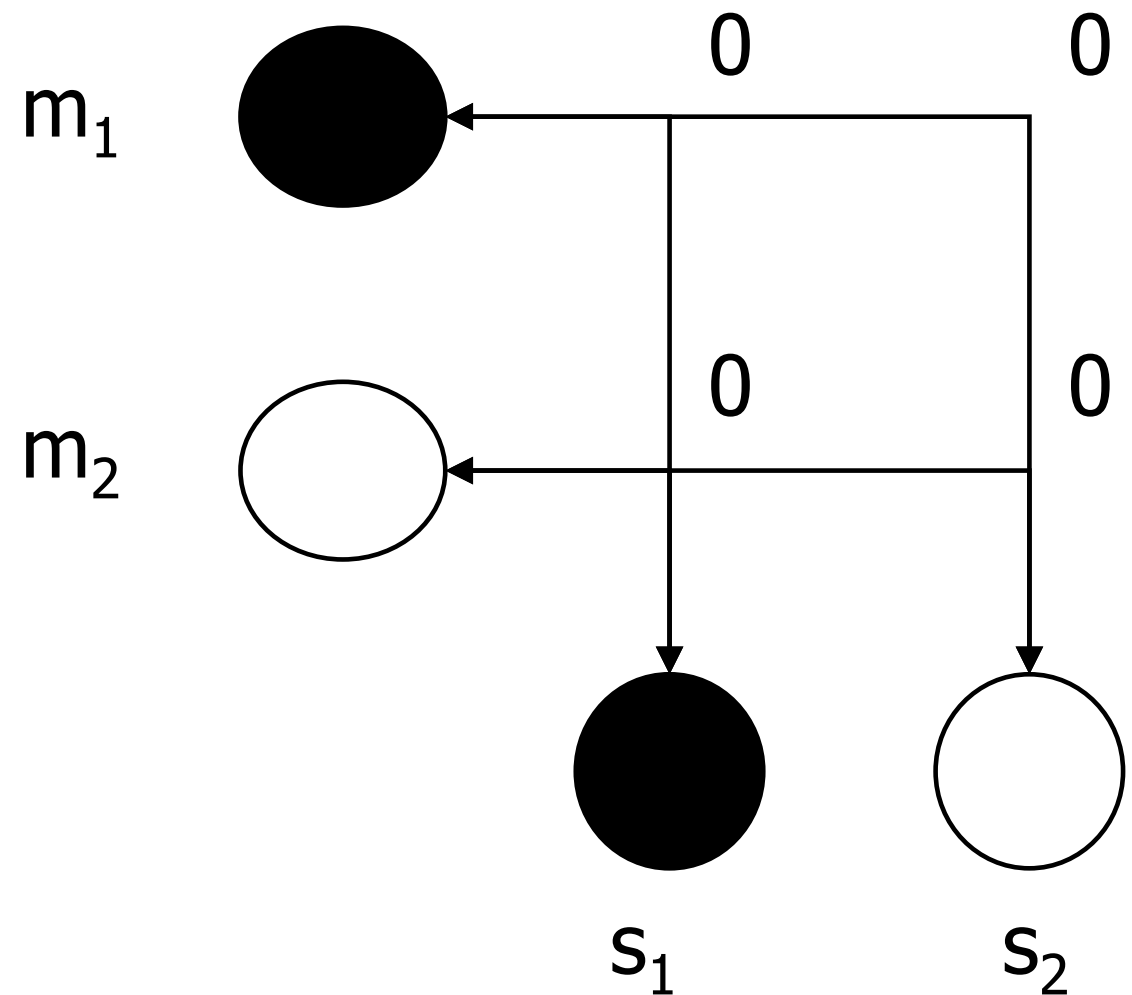
construction: $\alpha > \beta$ & $\delta > \gamma$

- What do these patterns mean?

Working out bias

- A constructor rule: $[+1, -1, -1, +1]$

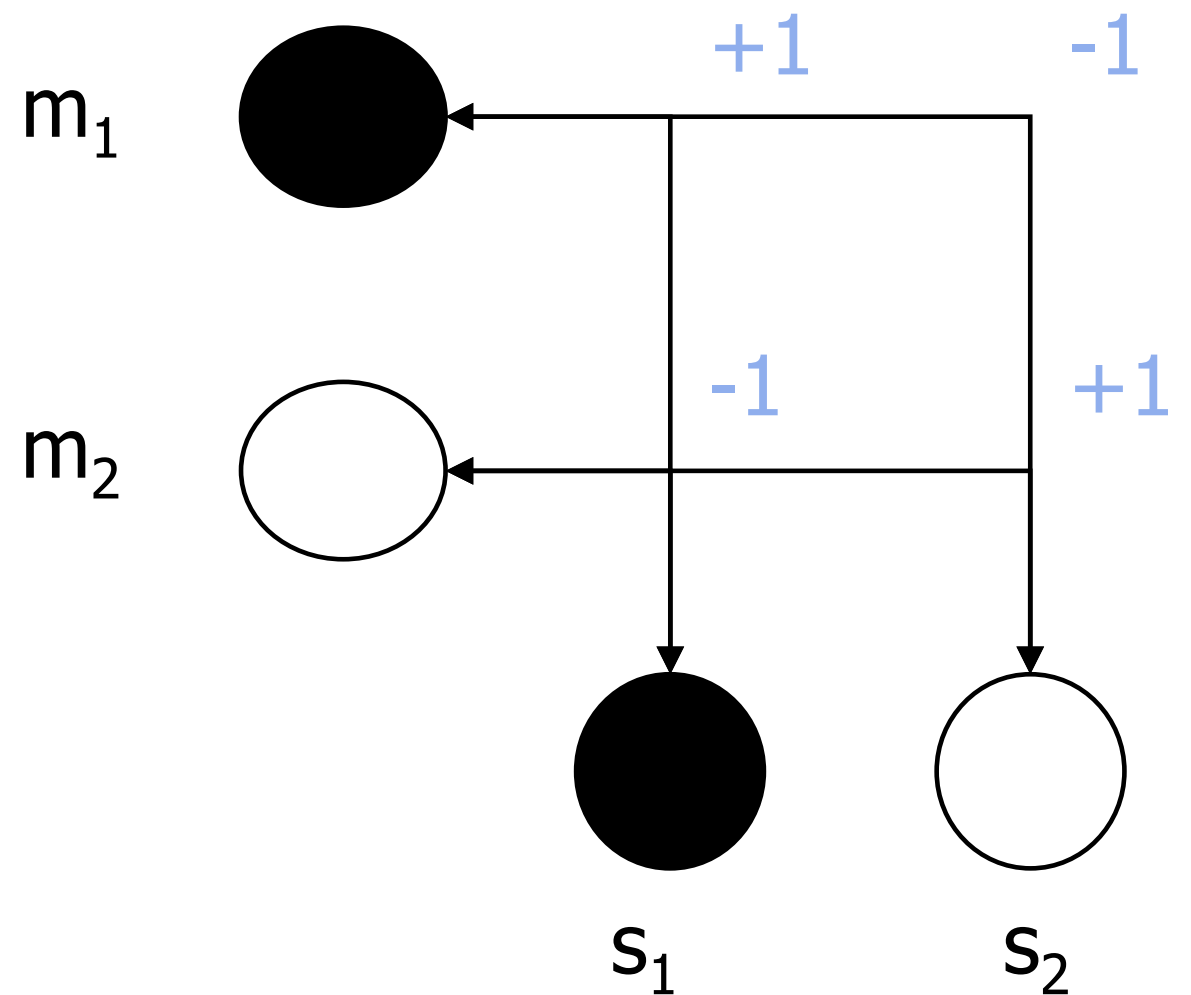
Observation:
 $m_1 \rightarrow s_1$



Working out bias

- A constructor rule: $[+1, -1, -1, +1]$

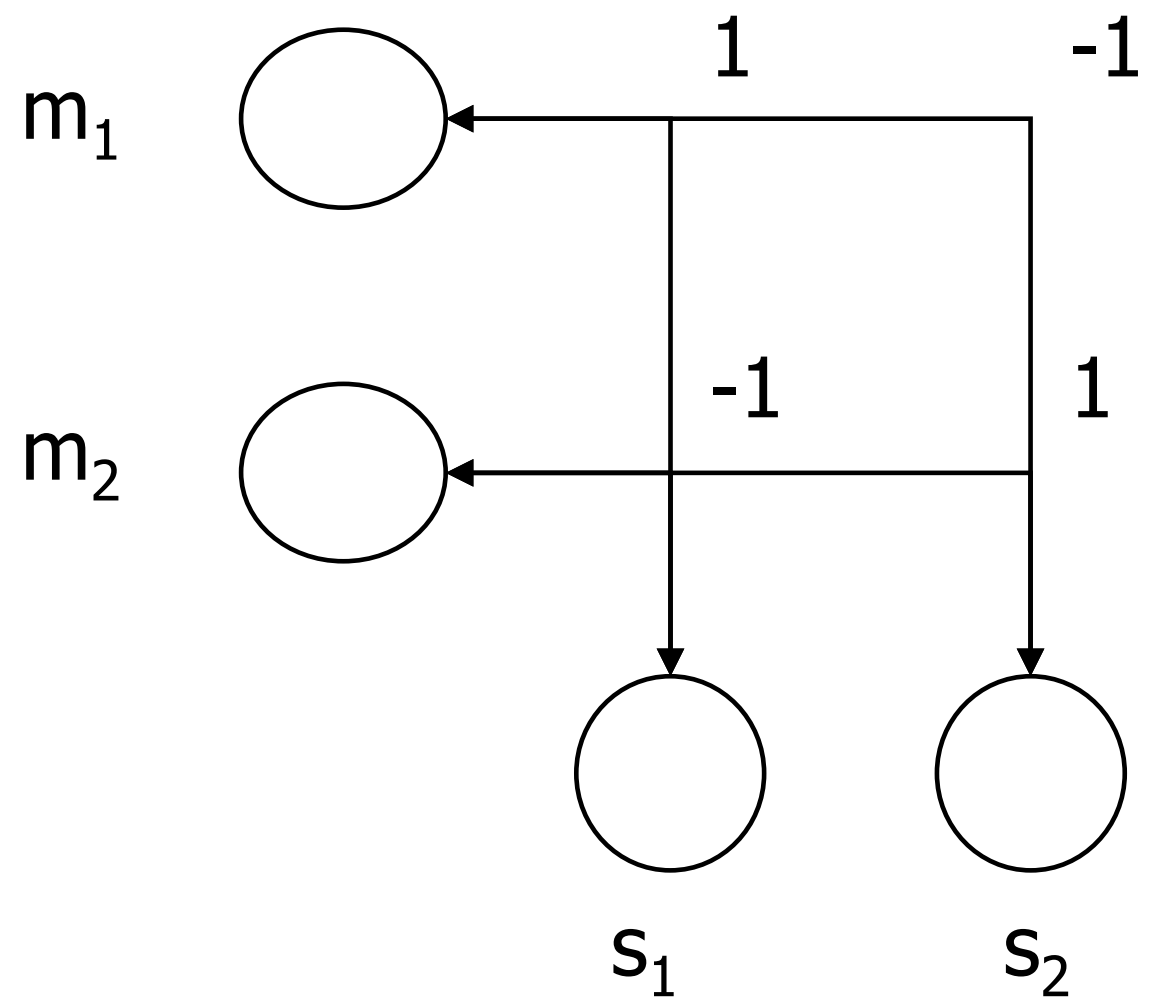
Observation:
 $m_1 \rightarrow s_1$



Working out bias

- A constructor rule: $[+1, -1, -1, +1]$

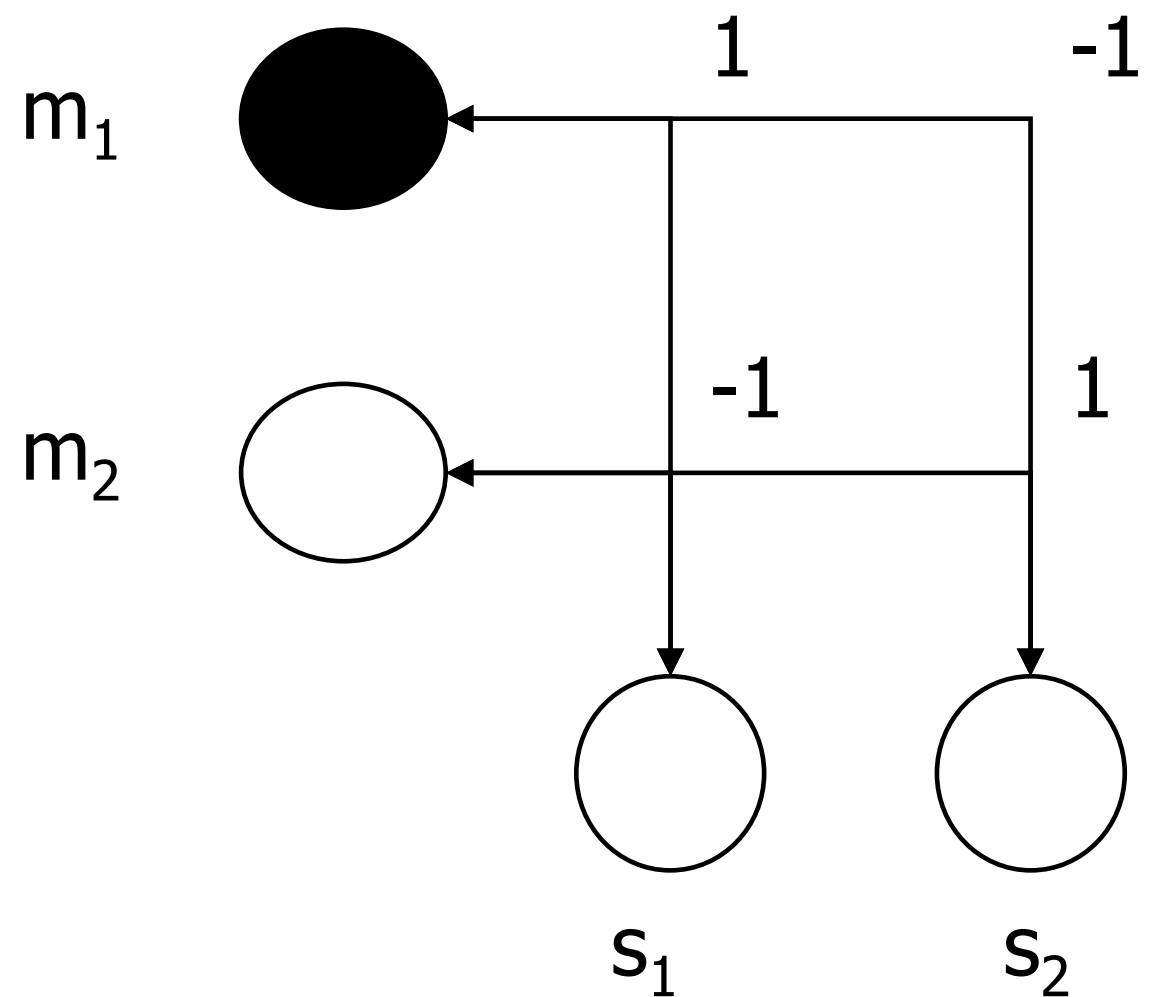
Observation:
 $m_1 \rightarrow s_1$



Working out bias

- A constructor rule: $[+1, -1, -1, +1]$

Production:
 $m_1 \rightarrow ?$



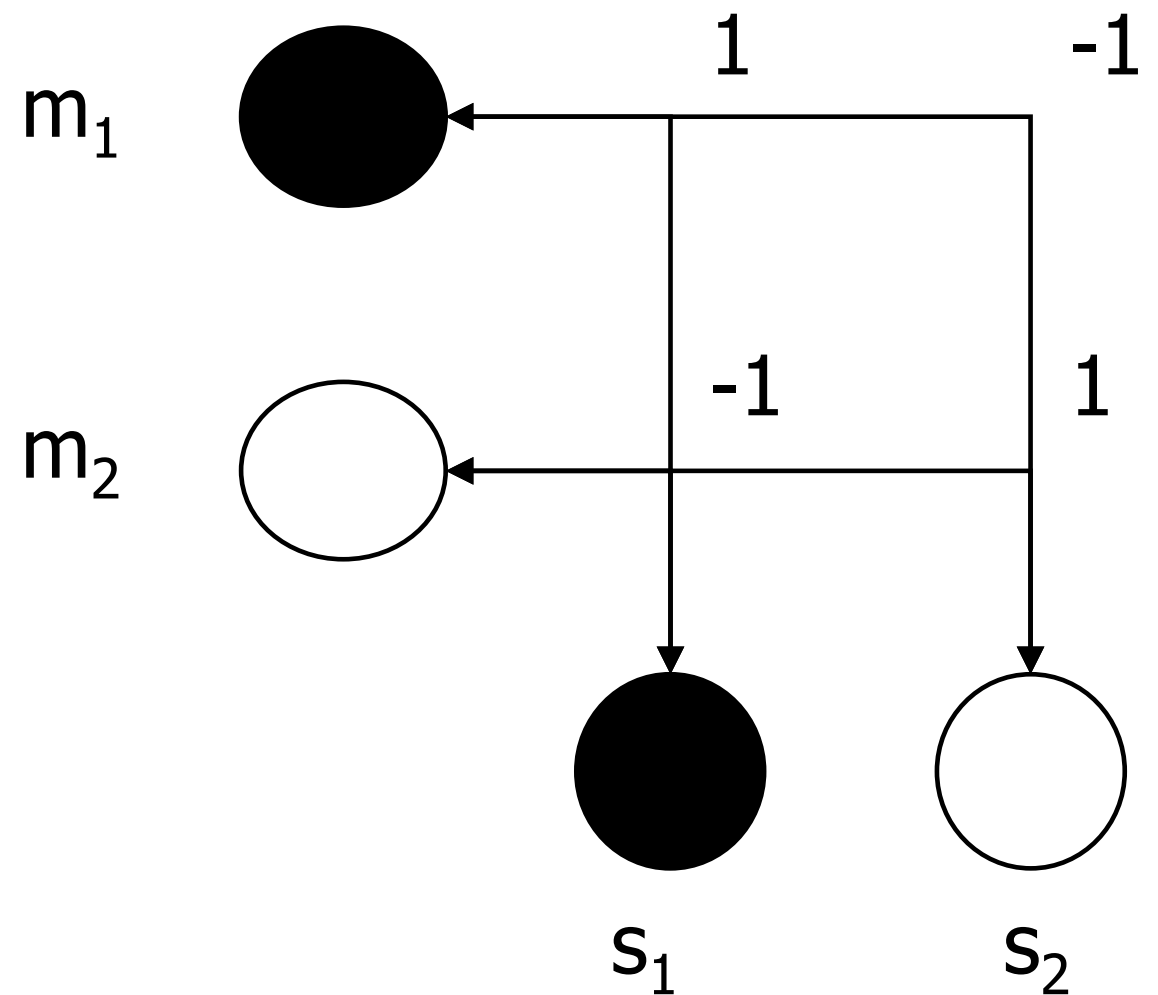
Working out bias

- A constructor rule: $[+1, -1, -1, +1]$

Production:

$m_1 \rightarrow s_1$

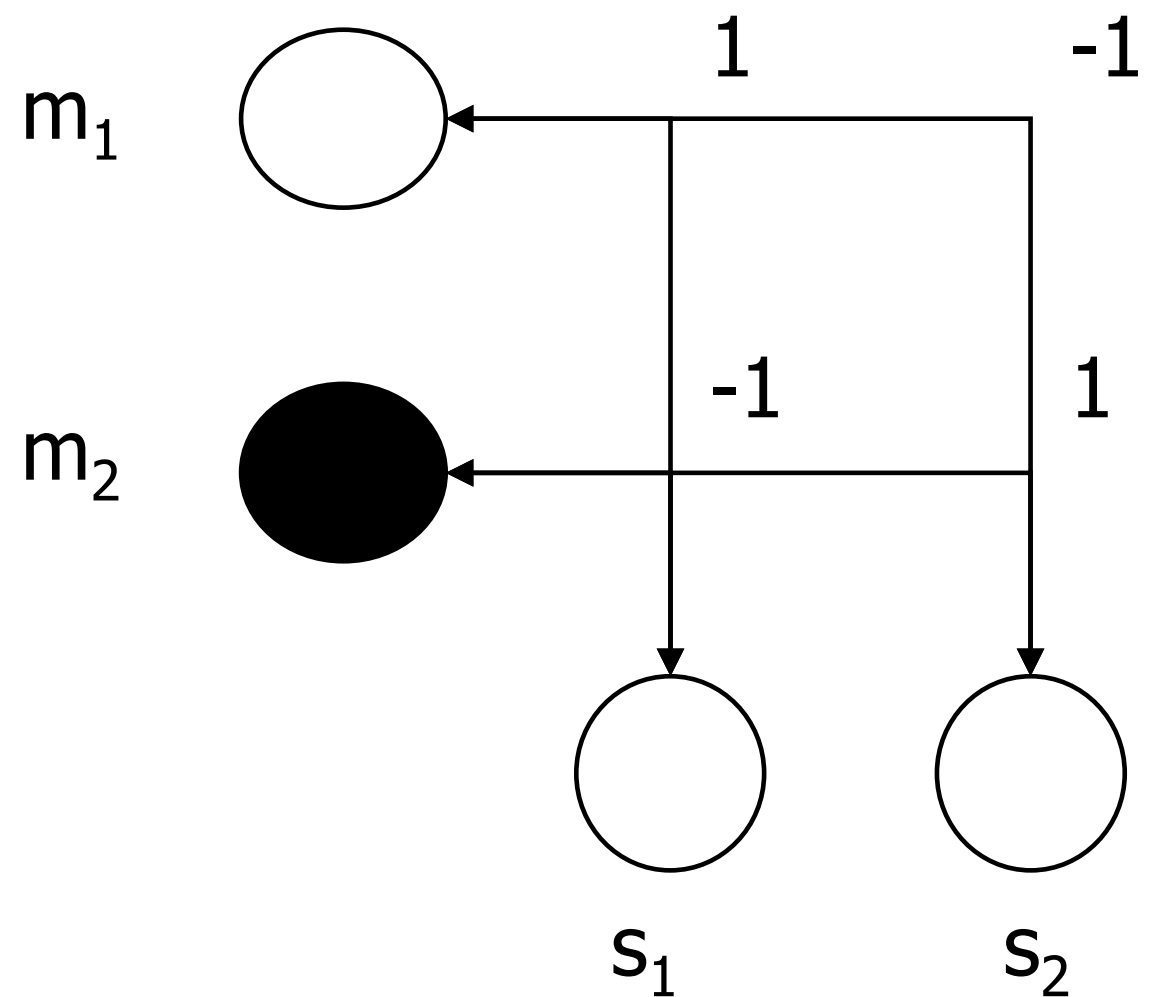
(*not* s_2)



Working out bias

- A constructor rule: $[+1, -1, -1, +1]$

Production:
 $m_2 \rightarrow ?$

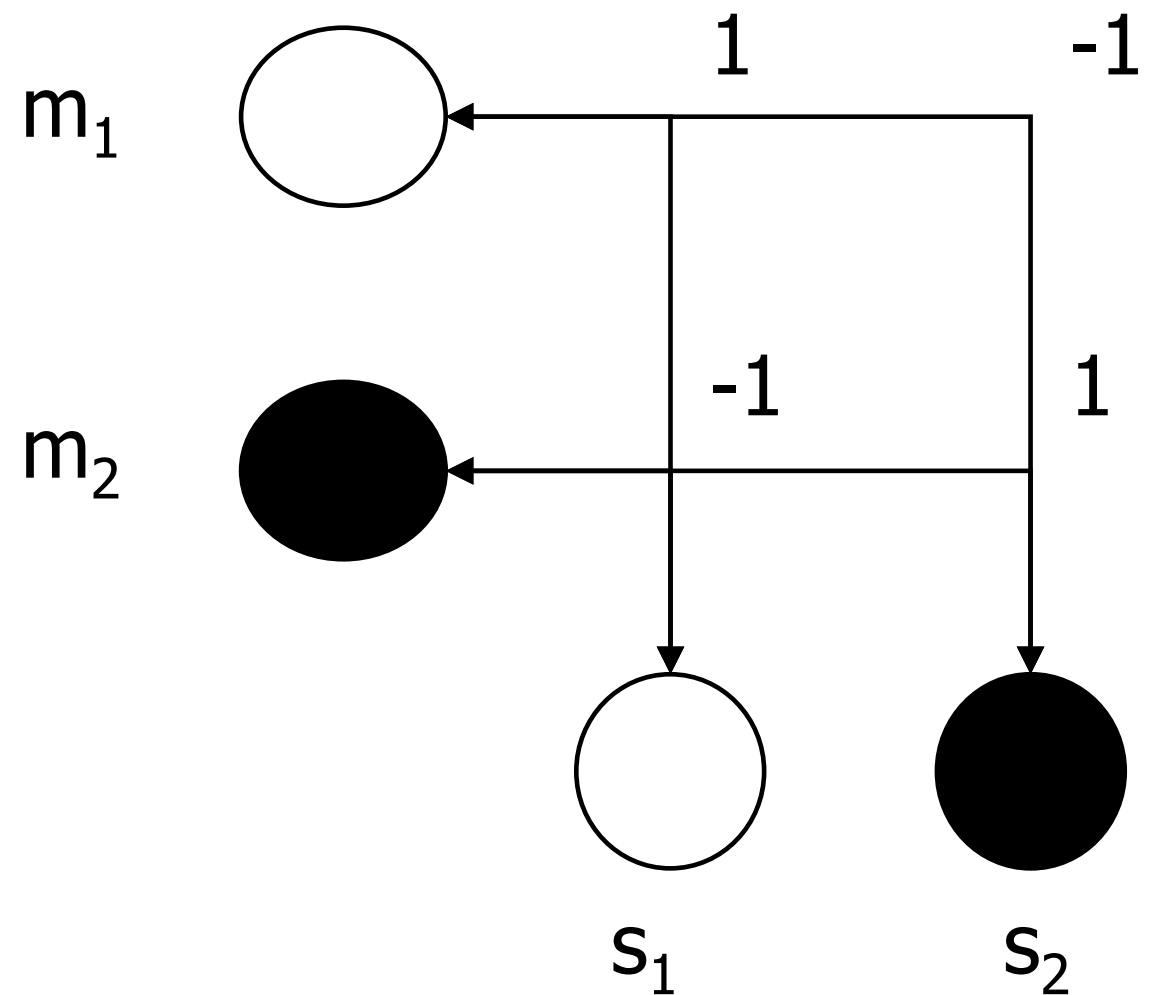


Working out bias

- A constructor rule: $[+1, -1, -1, +1]$

Production:

$m_2 \rightarrow s_2$
(*not* s_1)



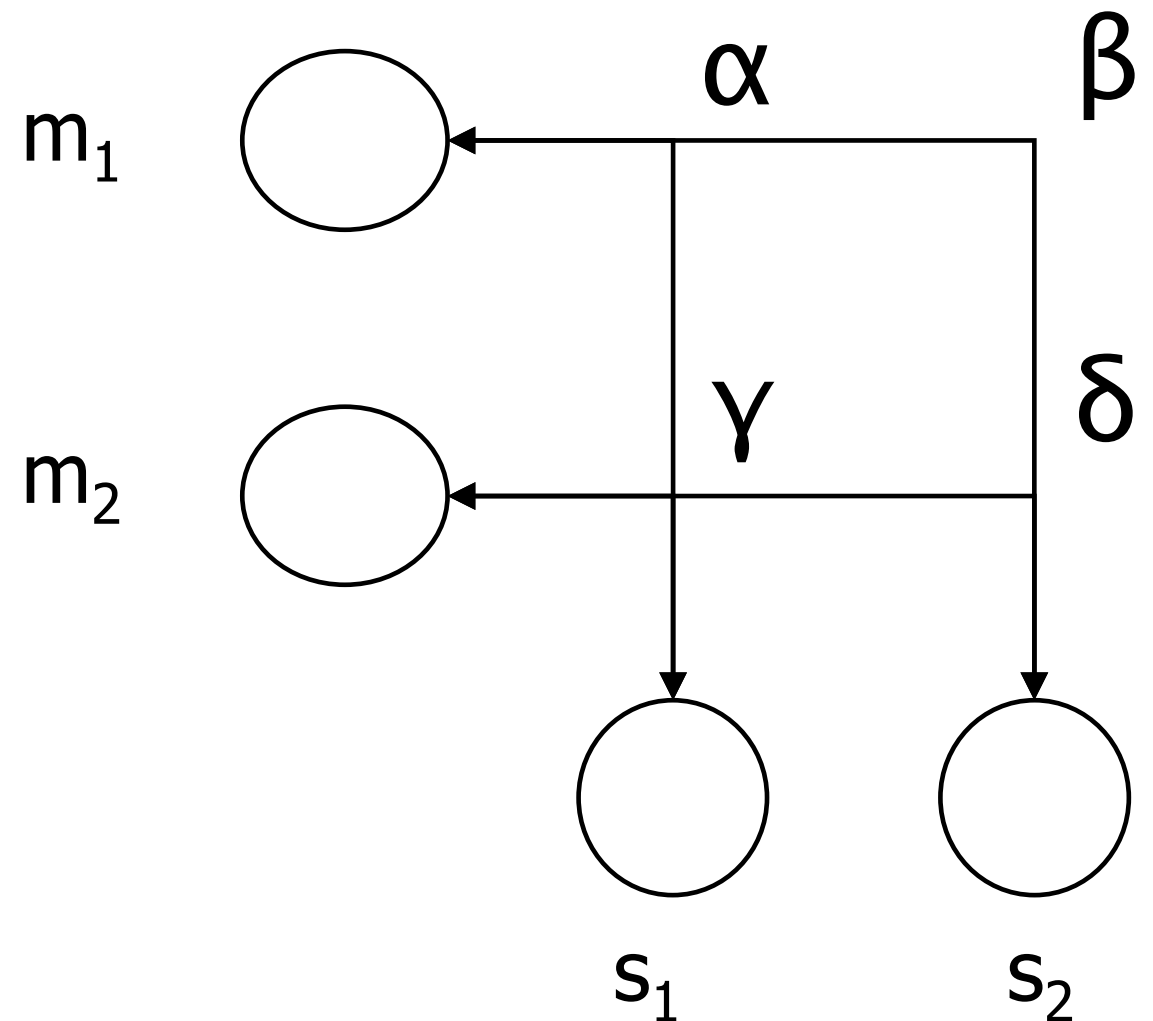
Working out bias

- Constructors in general: $\alpha > \beta$ & $\delta > \gamma$
After one exposure to $m_1 \rightarrow s_1$

Production:

$m_1 \rightarrow s_1$

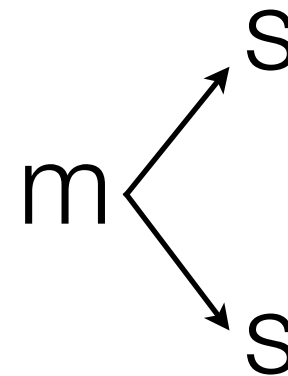
$m_2 \rightarrow s_2$



The constructor bias

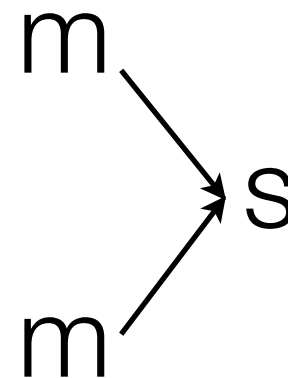
- Constructors don't like:
- One meaning to multiple signals

because $\alpha > \beta$
bias against synonymy



- Multiple meanings to one signal

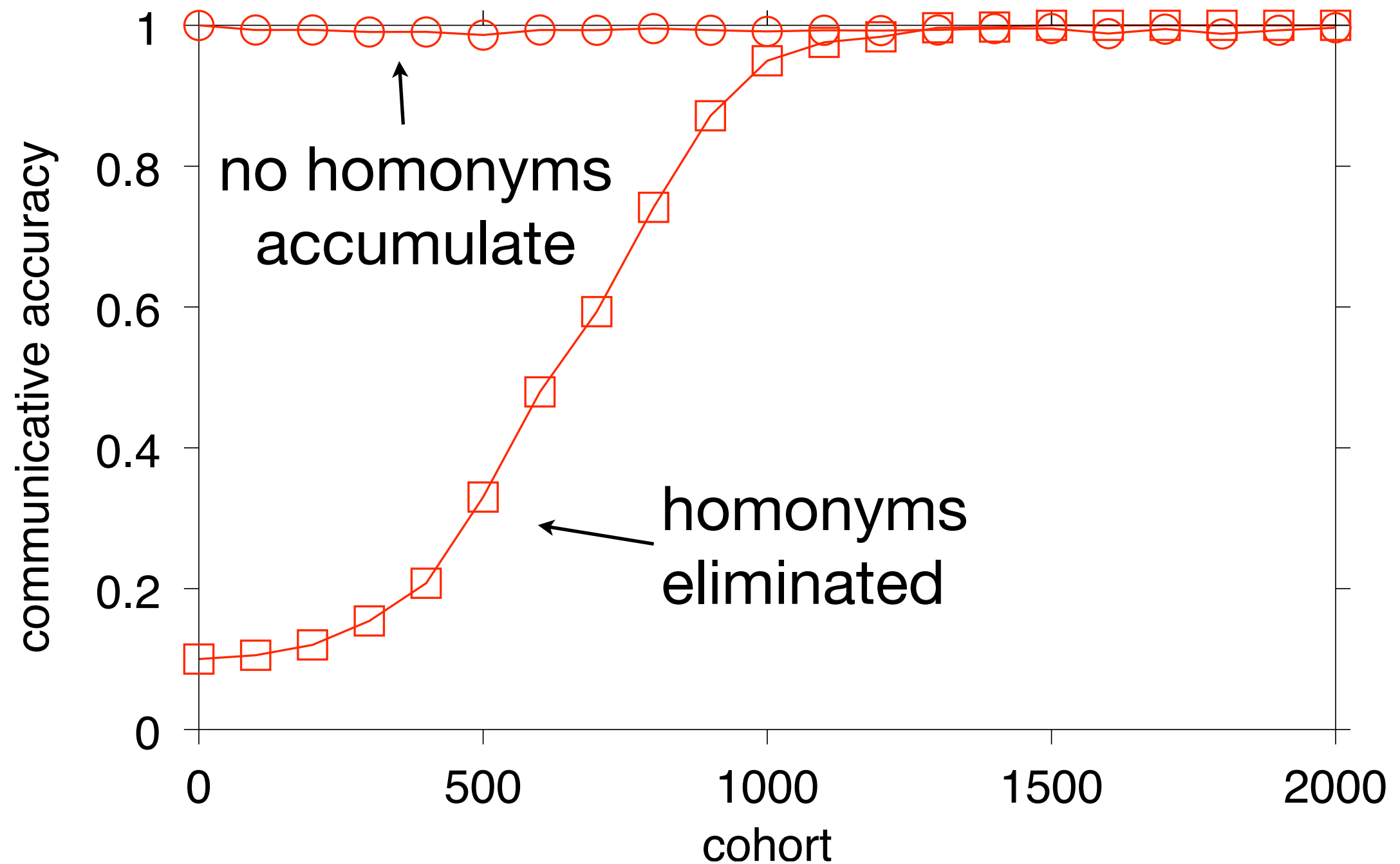
because $\delta > \gamma$
bias against homonymy



The constructor bias

- Constructors biased in favour of **one-to-one** mappings between meanings and signals
- Population's vocabulary changes over time to match this bias
- One-to-one systems happen to be optimal for communication

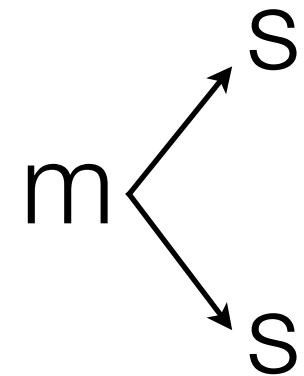
Constructor behaviour



The maintainer bias

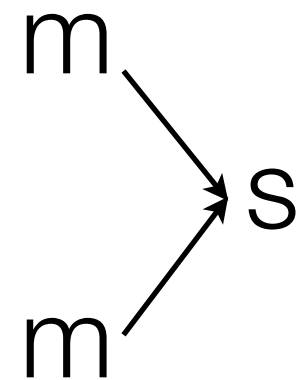
- Biased against synonymy

because $\alpha > \beta$

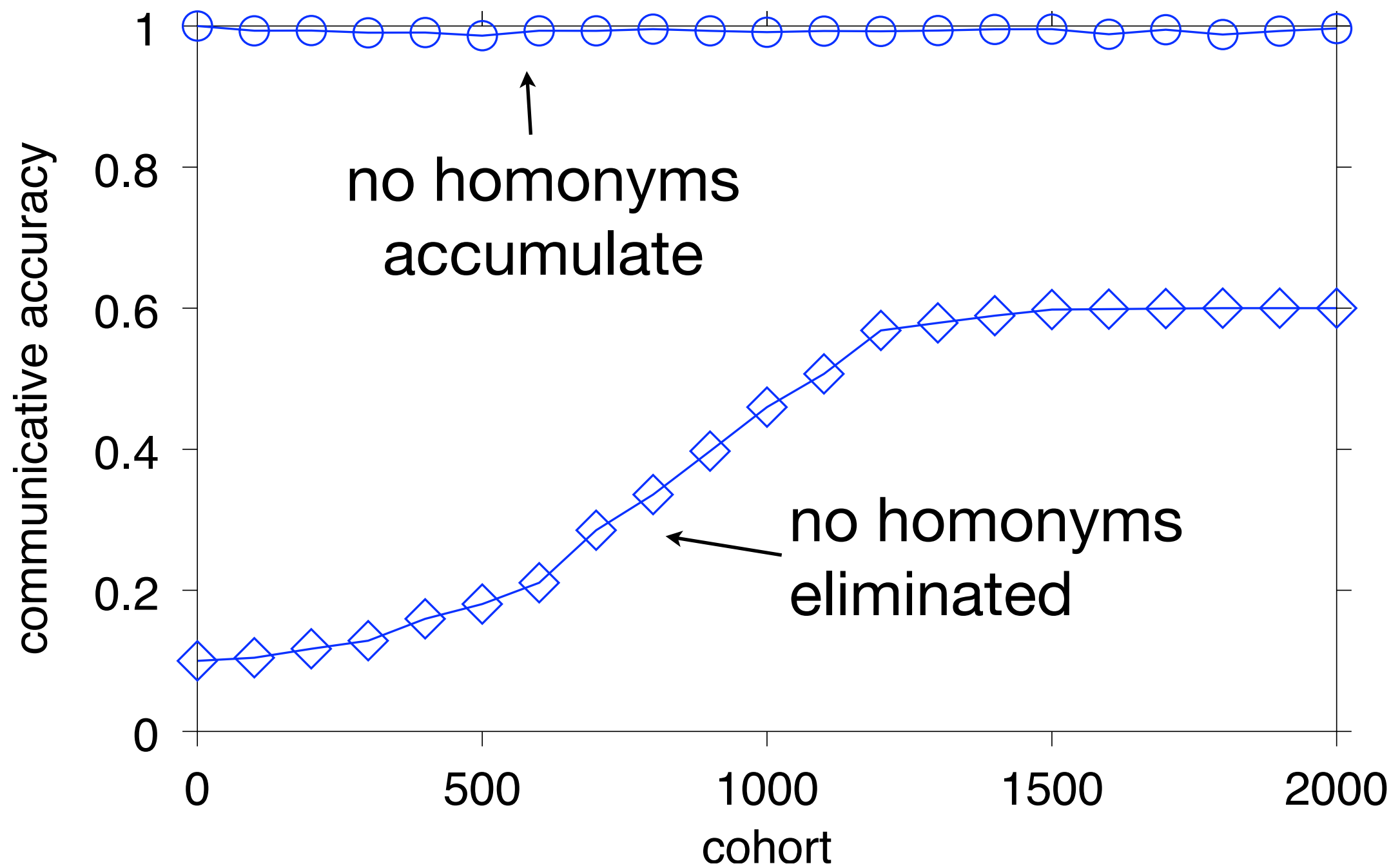


- Neutral with respect to homonymy

because $\delta = \gamma$



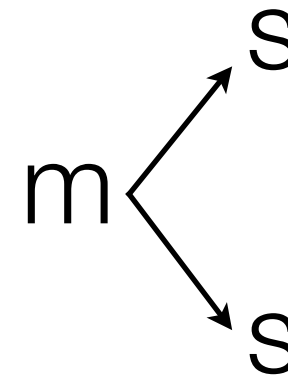
Maintainer behaviour



The learner bias (in most cases)

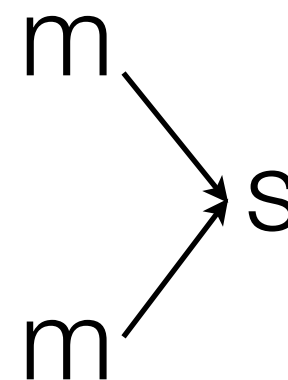
- Biased against synonymy

because $\alpha > \beta$

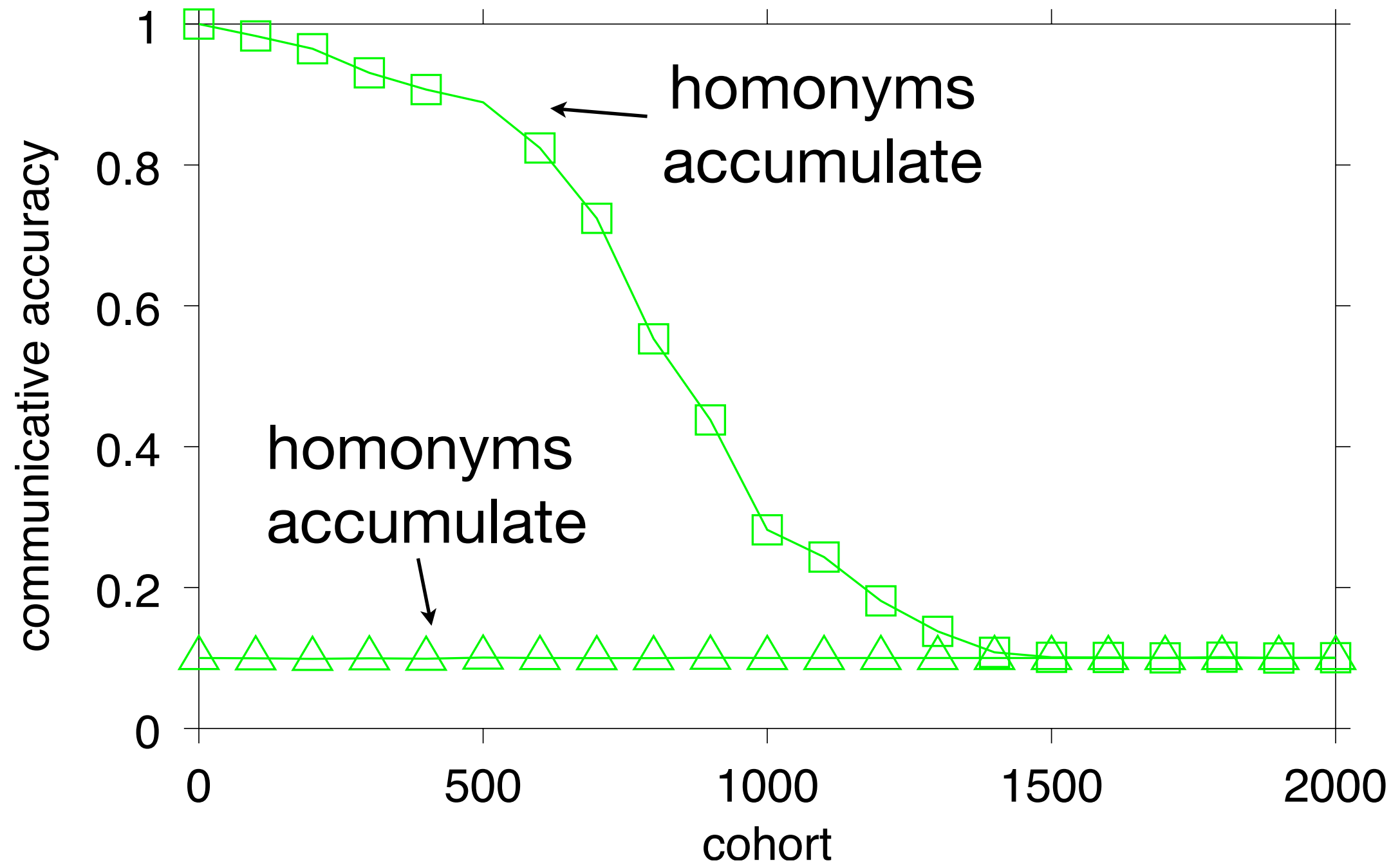


- Biased *in favour* of homonymy

because $\delta < \gamma$

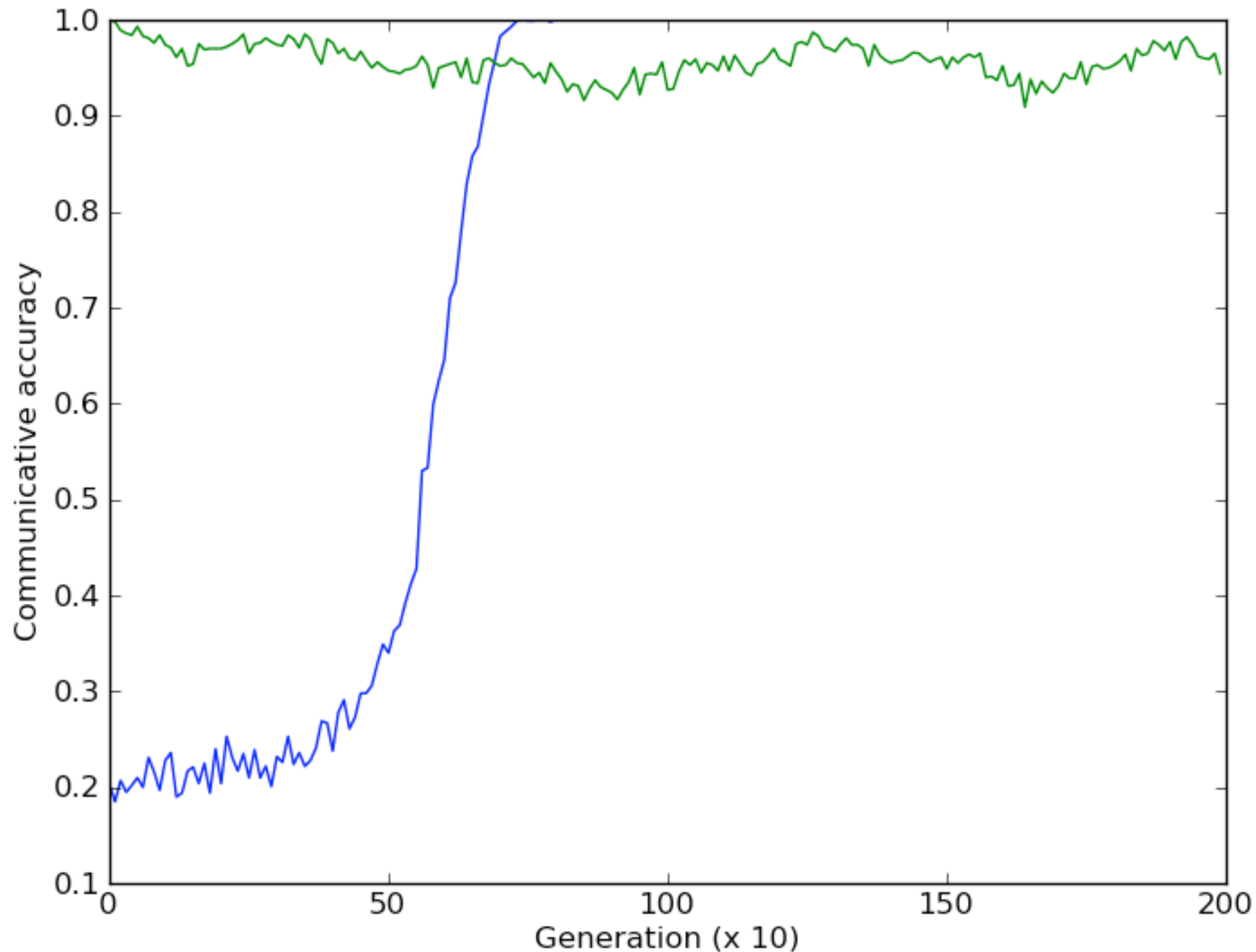


Learner behaviour



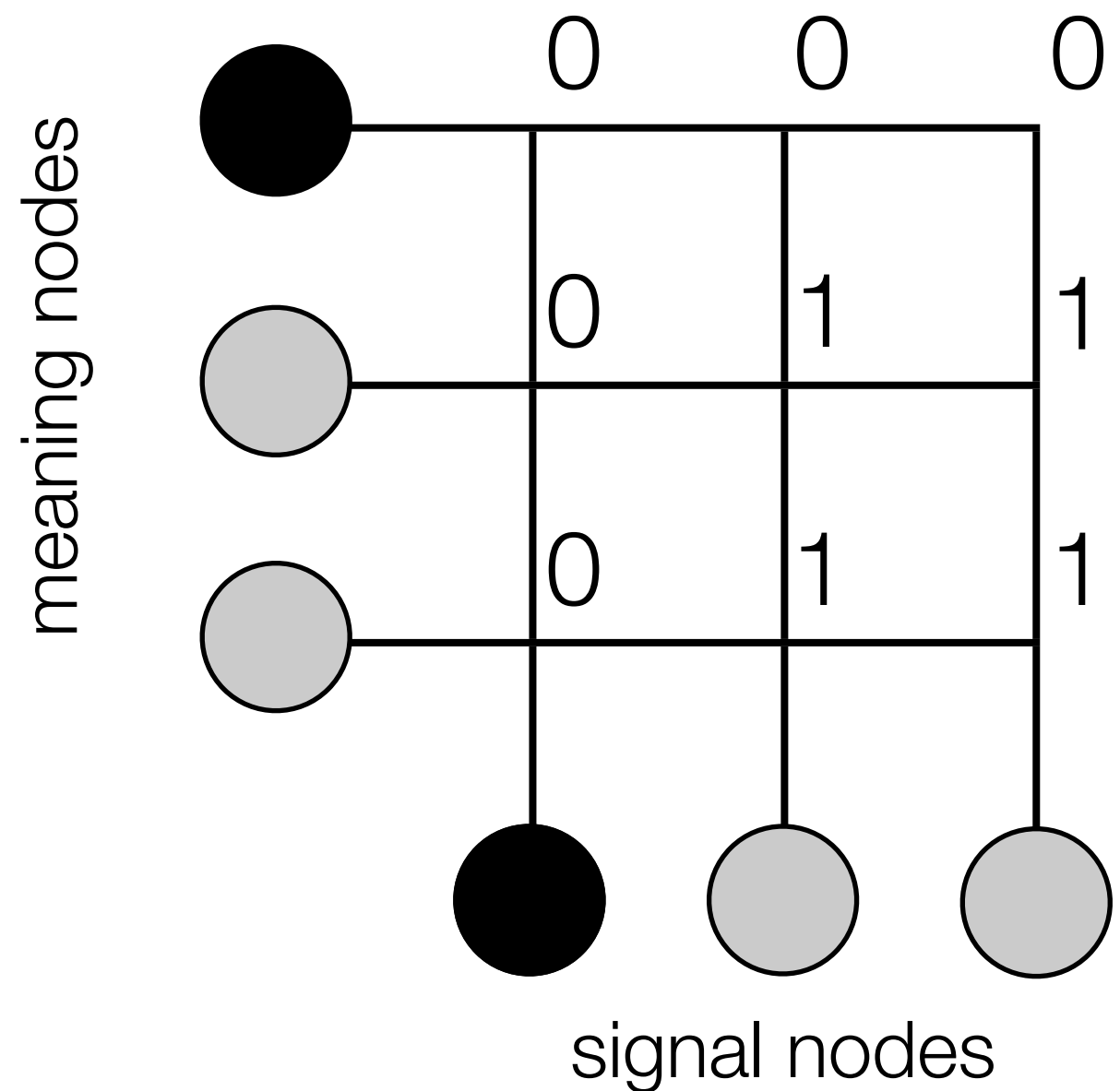
A problem (thanks to students on this course!)

- [0,0,0,1]: should be [+learner, -maintainer, -constructor]



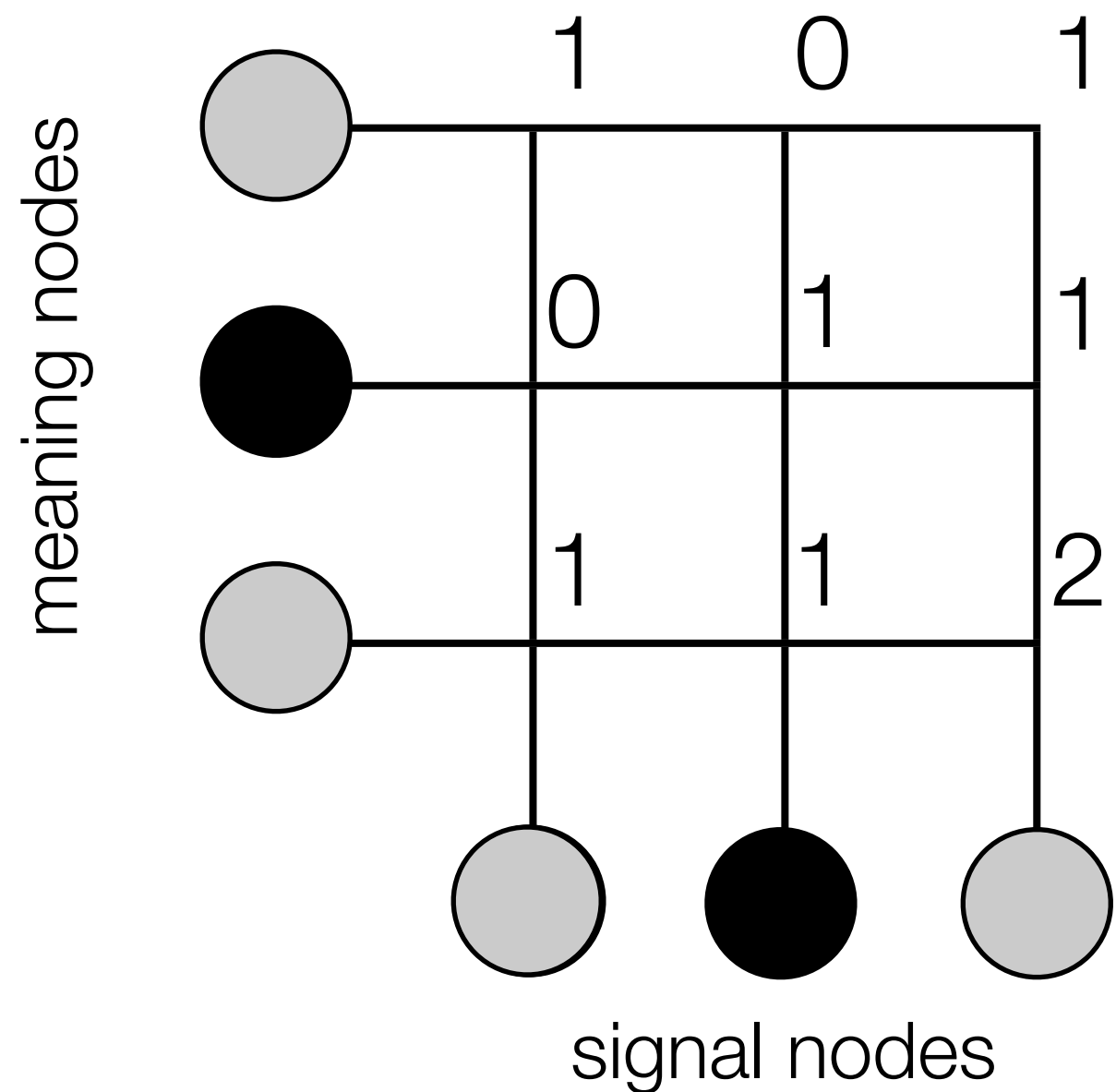
An anomalous rule: learning by co-non-occurrence

Observation:
 $m1 \leftrightarrow s1$



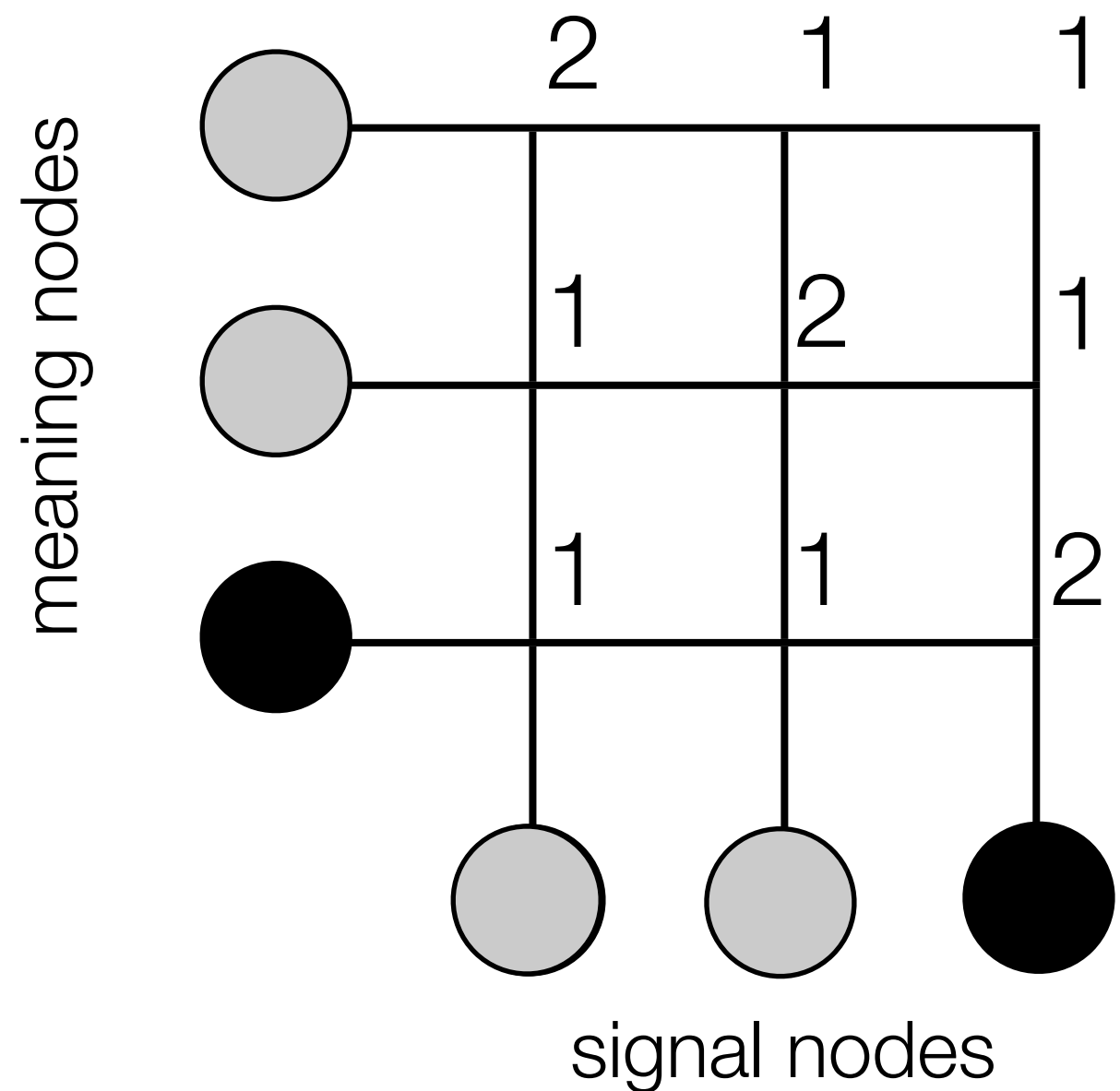
An anomalous rule: learning by co-non-occurrence

Observation:
 $m2 \leftrightarrow s2$



An anomalous rule: learning by co-non-occurrence

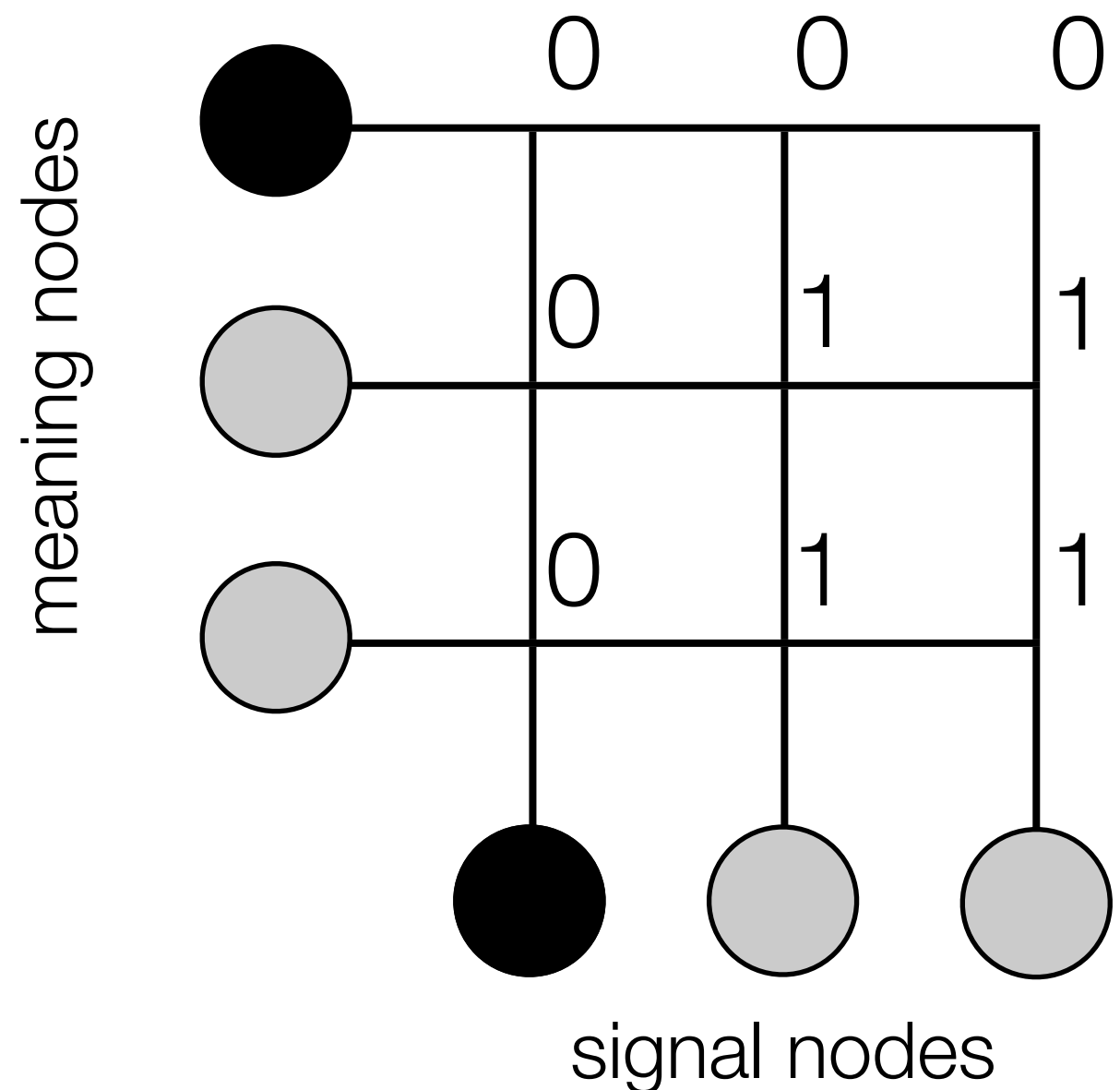
Observation:
 $m3 \leftrightarrow s3$



An anomalous rule: learning by co-non-occurrence

Observation:
 $m1 \leftrightarrow s1$

This looks like a 1-to-1
bias - that's why it
constructs and
maintains



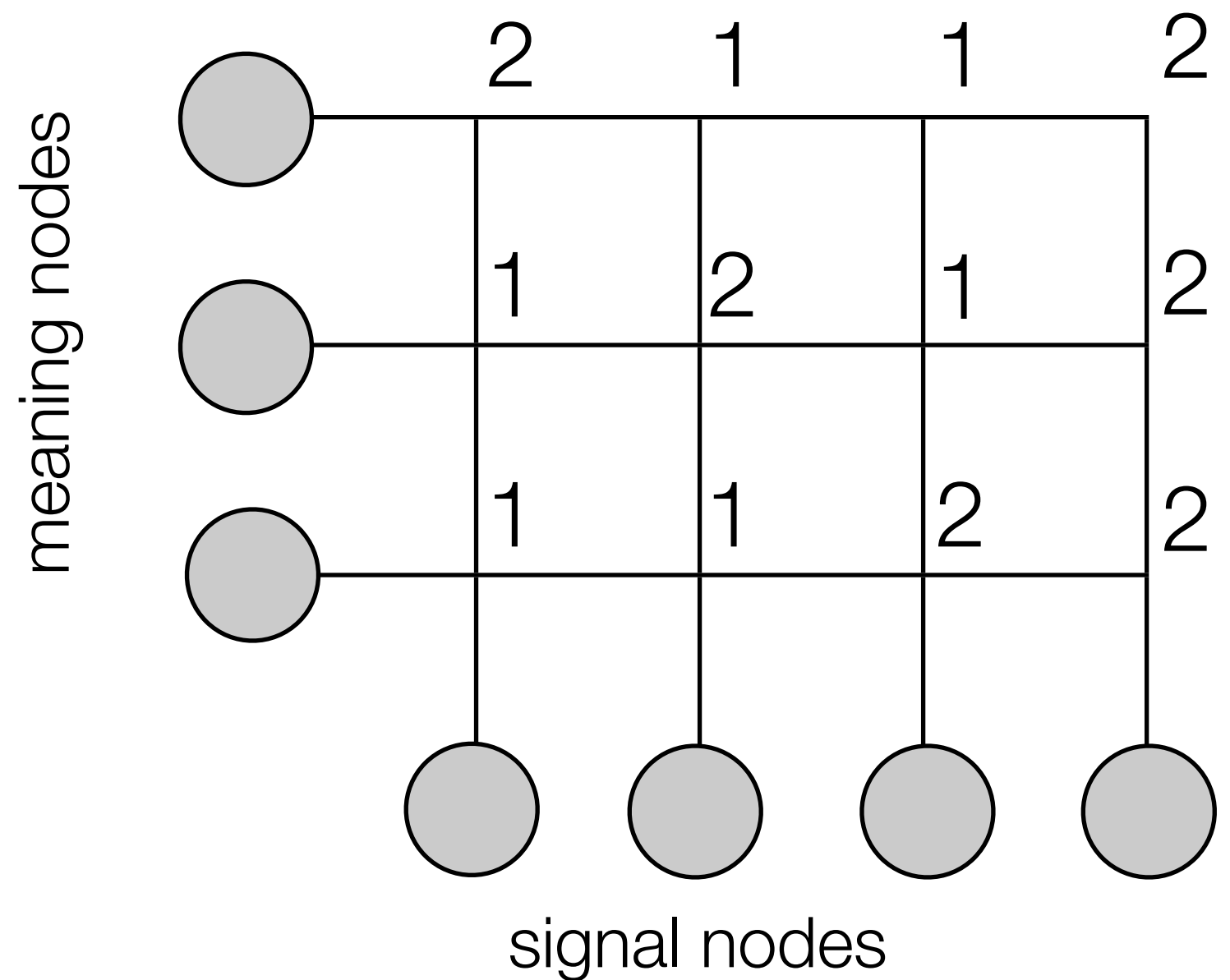
But ... adding more signals breaks it

Observations:

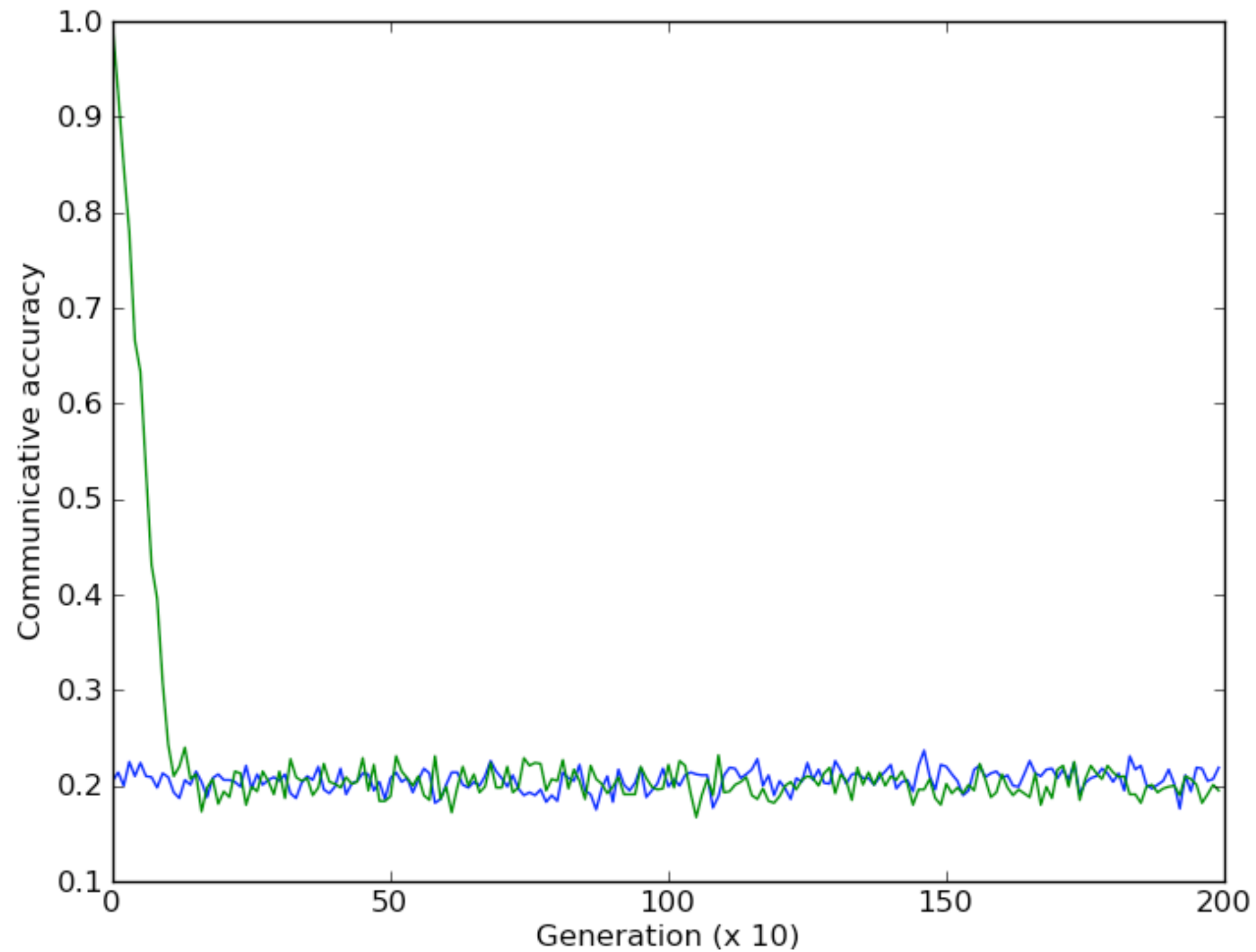
$m1 \leftrightarrow s1$

$m2 \leftrightarrow s2$

$m3 \leftrightarrow s3$



5 meanings, 20 signals



A modified acquisition criterion

For all learners: $\alpha + \delta > \beta + \gamma$

For all non-learners: $\alpha + \delta \leq \beta + \gamma$

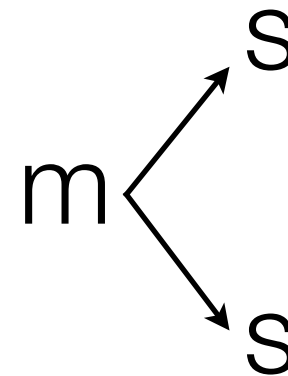
Additionally: if $|s| > |m|$, $\alpha > \beta$ required for acquisition

- Missed this in Smith (2002)
 - Slightly different implementation made anomalous behaviour less obvious
- Included in Smith (2004)

The constructor bias

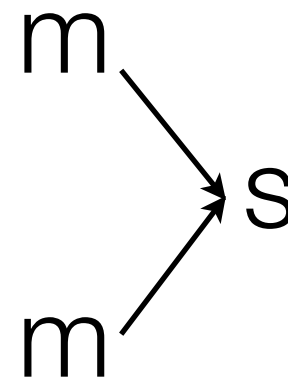
- Constructors don't like:
- One meaning to multiple signals

because $\alpha > \beta$
bias against synonymy



- Multiple meanings to one signal

because $\delta > \gamma$
bias against homonymy



What about real humans?

- Experiment on children's learning bias
Markman & Wachtel (1988) on synonymy



“Show me the fendle.”

- Children pick the unfamiliar object given an unfamiliar word

Synonymy bias

Before



banana



???

After (two possibilities)



banana
fendle

banana



fendle



Homonymy bias (Doherty 2004)

- “... at the zoo, they saw a strange tapir from Brazil. Hamish thought the tapir’s long nose looked funny”



“Which one is the tapir in this story?”

Homonymy bias (Doherty 2004)

- “... at the zoo, they saw a strange **cake** from Brazil. Hamish thought the **cake's** long nose looked funny”



“Which one is the **cake** in this story?”

Homonymy bias

Before

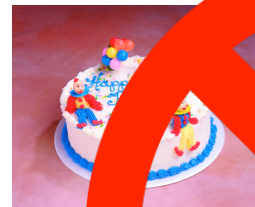


cake



???

After (two possibilities)



cake



cake



cake



???

Children's learning biases

- Children don't like:
 - synonymy
 - homonymy
- They have the same biases as constructors in our simple model
- Populations of constructors evolve optimal communication systems

A co-evolutionary hypothesis (Smith 2004)

Children's learning biases have evolved through natural selection, because they're good for communication.

- Examine this idea using our model
- Two central assumptions:
 - Weight update rule is given by a genotype
 - Better communicators breed more

Invasion of the mutants

- Smith (2004) plays **constructors**, **maintainers**, and **learners** off against each other
- Create a population mainly made up of one type, but with a small number of another type (the mutant)
- Agents inherit both the communication system (by cultural transmission), and their learning strategy (by genetic transmission)
- Both culture and biology evolve
- If selection is based on communicative success, which mutants will invade?

Surprising result: evolution is hard

- Constructors don't often invade, *even though it would increase the fitness of the population if they did*
- Two problems:
 - Need a lot of mutants before they start to have a good effect on the population's language...
 - ...and even then, there's a time-delay before the good language evolves culturally.
- Speculative conclusion: human learning biases *haven't* evolved only for communication.

Summary

- Smith (2002, 2004) look in detail at how learning bias can give us (or fail to give us) language
- Brings together 3 complex processes in one model:
 - Learning
 - Cultural transmission
 - Biological evolution
- Highlights the crucial importance of the second of these three
- BUT... language model is extremely simple. Next we'll look at making acquisition of meanings more realistic, and then we'll have a look at models of the evolution of more complex signals (i.e. syntax)

Reading

- Smith, K. (2004) The evolution of vocabulary. *Journal of Theoretical Biology*, 228, 127–142

Extends the model in the previous paper to look at evolution of bias by examining invasion of mutants.