Notes 2.0 on presentation of model for co-evolution

Thomas Brochhagen

Motivation

We would like to present our model in a cleaner and more intuitive way. As suggested by Michael, one way to do so is to focus first on a reduced type space to illustrate how the model works. As a reminder:

$$L_{\text{bound}} = \begin{array}{ccc} & m_{\text{some}} & m_{\text{all}} \\ s_{\exists\neg\forall} & 1 & 0 \\ s_{\forall} & 0 & 1 \end{array} \qquad L_{\text{lack}} = \begin{array}{ccc} & m_{\text{some}} & m_{\text{all}} \\ s_{\exists\neg\forall} & 1 & 0 \\ s_{\forall} & 1 & 1 \end{array}$$

These two lexica are paired with either level-0 *literal* behavior, or level-1 *pragmatic* behavior. Learnability-wise, if there is a bias for simpler semantics (no-upper bound), then the two types with L_{lack} are favored *a priori*. Furthermore, both L_{bound} users are indiscernible when it comes to learning; they only differ in their receiver behavior. Utility-wise, literal L_{bound} fares slightly better than pragmatic L_{bound}. If λ is sufficiently high, then the difference between pragmatic L_{lack} and both L_{bound} types is almost negligible. We want to visualize how (co-)evolution plays out by comparing four populations of two types each, where each pair of types differs only in (i) signaling behavior, or (ii) lexica:

1. literal L_{bound} vs. literal L_{lack};
2. pragmatic L_{bound} vs. pragmatic L_{lack};
3. literal L_{bound} vs. pragmatic L_{bound};
4. literal L_{lack} vs. pragmatic L_{lack}.

The other relevant parameters of the model are: soft-max parameter λ, posterior parameter l, sequence length k, and bias parameter c (detailed below).

Learning bias

In the paper, the learners' inductive bias is a function of the complexity of the formulae that types lexicalize. Introducing our measure and applying it only to this small type space is probably not the way to go. For illustratory purposes, we may instead assume that there's an abstract parameter (which we later make concrete using our measure) $c \geq 1$ that regulates how strongly L_{lack} is preferred over L_{bound}. In what follows, I assumed this to be $P(L_i) \propto c$ if L_i is L_{lack} and otherwise 1. For instance, if we have a type space with τ_i as literal L_{bound} and τ_j as pragmatic L_{lack}, then $P(\tau_i) \approx .33$ for $c = 2$. This is just for illustration, of course. We can change this at will. As far as I can see, the reduced type spaced behaves in a very predictable manner.

General remarks 2.0

There are three remarks to be made about this update.

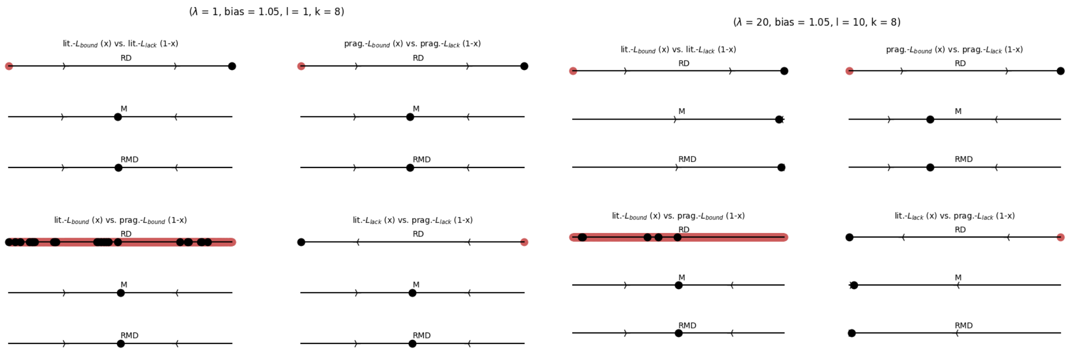
First, the phase plots are a lot cleaner and all show stationary points. These were either calculated, or estimated as being in the middle of two points that changed in opposite directions. As before, the RD competition between literal and pragmatic L_{bound} is messy because there are many points for which $RD(x) = x$. On this front, I just changed white circles with black edges to red circles without edges for better visibility when many circles overlap.

Second, when it comes to the direct competition of only two types on a line (phase plots), M and RMD still produce pretty similar outcomes. I tested a range of learning bias values and larger k but there's not much change. This is not really troublesome (it being a consequence of there only being two types in a population), but I think this may disqualify the phase plots as candidates for a good illustration of the model's dynamics.

Third, I tried my hand at the 2D quiver plots with all types, not only at the edges but also at the center of the unit square. As Michael suggested, (x, y)-pairs in the square give the following proportions of types:

- $lit.\ L_{\text{bound}}$: $(1 - x) * (1 - y)$
- $lit.\ L_{\text{lack}}$: $(1 - x) * y$
- $prag.\ L_{\text{bound}}$: $x * (1 - y)$
- $prag.\ L_{\text{lack}}$: $x * y$

Overall, I think these plots look good. Important, to me, is that (i) M and RMD are clearly distinguishable and that (ii) it's clear that RD with lowish λ leads to L_{bound}. Another plus is that it's easy to explain how this plot came about. On the negative side, I'm



(a)

(b)

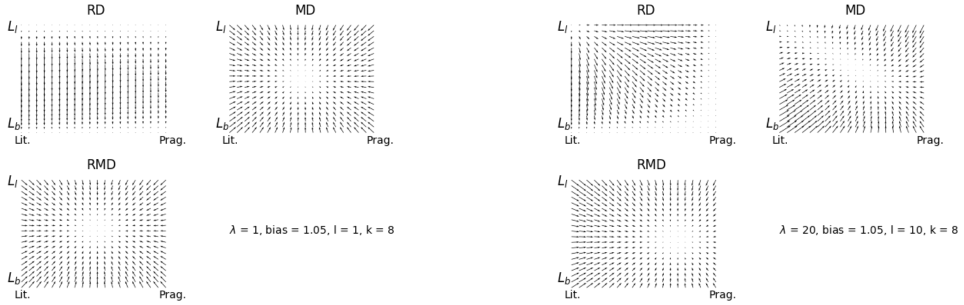
Figure 1: Approximate phase-portraits for (a) $\lambda = 1, c = 1.05, l = 1, k = 8$ and (b) $\lambda = 20, c = 1.05, l = 10, k = 8$. Attractors are in black. Non-attractor points are in red for better visibility for lit. L_{bound} vs prag. L_{bound}, which has many.

still a bit unhappy about having to assume that prag. and L_{lack} are independent, but this is clearly better than just plotting the edges. The one important I am missing from these plots is clear movement on the lower edge, from $(1, 0)$ to $(0, 0)$, which would indicate movement from prag. L_{bound} to lit. L_{bound}, at least for RD-only. This would help us in driving the message home that pragmatic recruitment is not functionally advantageous for all types. I think the problem right now is that movement on this edge progresses in very small steps (if at all; cf. phase plots). These small steps disappear once we scale them with the bigger change that happens elsewhere in the square. Reinterpreting (x, y) would not help on the edge itself, but maybe taking average trajectories would at least make points close to the edge move toward lit. L_{lack} rather than toward its pragmatic counterpart.

Phase portraits

As before, Figure 1 shows 8 (approximate) portraits for applications of replicator steps only, mutator steps only, and replicator-mutator steps on our four type-pairings, for two parameter constellations (c is now 1.05 instead of 2). For instance, the upper-left-most portrait shows the dynamic for a type space consisting of literal L_{bound} (x) and literal L_{lack} $(1 - x)$. We can see that, as long as there are some L_{bound} types in the population, the population progresses toward the extreme where only this type survives.

Beyond having cleaner portraits, my take on these is as before: They are nice because it's easy to see how the dynamics play out but they do not highlight difference between $M(x)$ and $M(RD(x))$ very well. The most pronounced differences I have



(a)

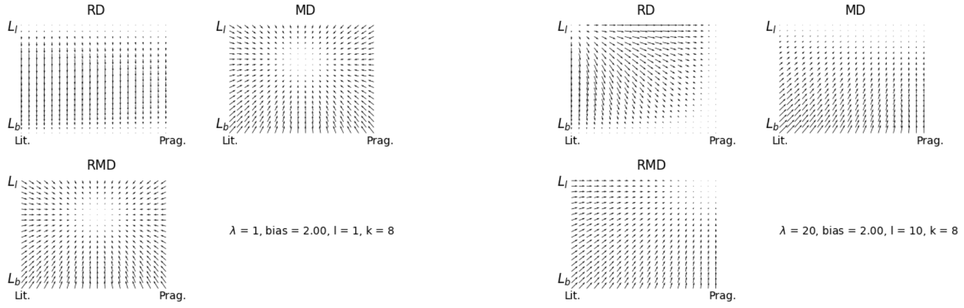
(b)

Figure 2: Evolutionary competition between pairs of types with (a) $\lambda = 1, c = 1.05, l = 1, k = 8$ and (b) $\lambda = 20, c = 1.05, l = 10, k = 8$. Coordinates (x, y) show (proportion of pragmatic types, proportion of L_{lack} types).

found are when λ is high, c low, and learners tend to sample from the posterior posterior sampling (e.g.; $\lambda = 20$, $c = 1.05$, $l = 1$). However, that's a rather esoteric parameter constellation to illustrate the dynamics with.

Dynamics

Figure 2 ($c = 1.05$) and Figure 3 ($c = 2$) illustrate the new filled unit squares. The main news are that we got rid of the ambiguity for (x, y) when we move away from the edges (see above). Taking stock: the good is that (i) the contrast between RD, M, and RMD is clear, (ii) we can see that lit. L_{lack} is functionally disadvantageous and that its pragmatic counterpart is so as well if λ is low, (iii) that a functional advantage *and* a learnability advantage will select for pragmatic L_{lack} but not one of these on its own. The bad is that we don't see the slight functional disadvantage that prag. L_{bound} suffers from when compared to its literal counterpart.



(a)

(b)

Figure 3: Evolutionary competition between pairs of types with (a) $\lambda = 1, c = 2, l = 1, k = 8$ and (b) $\lambda = 20, c = 2, l = 10, k = 8$. Coordinates (x, y) show (proportion of pragmatic types, proportion of L_{lack} types).