# Base games

Bas Cornelissen

August 9, 2017

This appendix mathematically derives the implicit biases in the additive naming game and presents some further analyses of the parameters of the game.

Implicit biases in the additive naming game

The additive base game strongly favours the use of the highest base. To quantify that bias, we ask what the probability is that an agent, without any past experience, will use a certain base in the next round. 'Without past experience' is important, since we are interested in the *bias*, similar to the Bayesian naming game. The problem, I think, becomes much more intuitive if you keep the following diagram in mind:
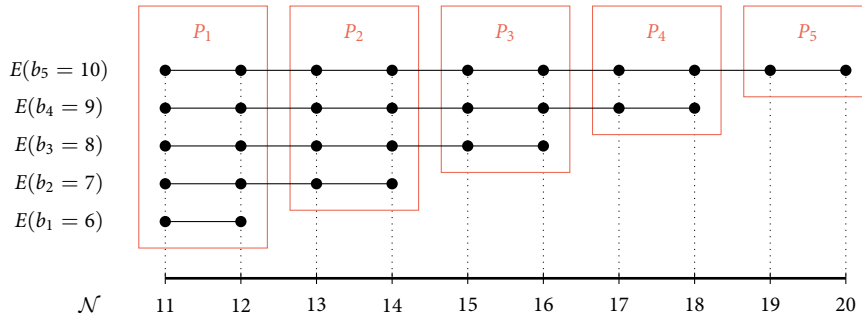
The diagram illustrates the decimal case $B = 10$, so $K = 5$. Recall that $E(b)$ is the set of all numbers n in the domain that are expressible with base b; these sets are shown as horizontal lines with dots at every number n. We first compute the probability $p(n \in E(b_j))$ that we pick a number expressible by the j'th base, and then the probability that base b_j is selected *given* that the number is expressible. From the diagram one directly sees that

$$p\big(n \in E(b_j) \big) = \frac{|E(b_j)|}{B} = \frac{j}{K},$$ (1)

which is the relative length of the black line corresponding to $E(b_j)$. Next, $p(b_j \mid n \in E(x_j))$ has to take into account that other bases might also express n. Given that $n \in$

$E(b_j)$ there are j *equally likely* 'parts' $P_1, \ldots, P_j$ that n could have been in. The parts correspond the orange boxes. Now it is easily seen that the numbers n in box P_j can be expressed by $K - j + 1$ different bases. In the diagram, the numbers 13 and 14 are in part P_2 and can be expressed by $5 - 2 + 1 = 4$ different bases. But that means that

$$p\big(\, b_j \mid n \in E(b_j)\,\big) = \sum_{i=1}^{j} p(n \in P_i) \cdot p(b_j \mid n \in P_i) \tag{2}$$

$$= \frac{1}{j} \cdot \sum_{i=1}^{j} \frac{1}{K - i + 1} \tag{3}$$

$$= \frac{1}{j} \cdot \Big(\sum_{i=1}^{K} \frac{1}{i} - \sum_{i=1}^{K-j} \frac{1}{i} \Big) \tag{4}$$

$$= \frac{1}{j} \big(H_K - H_{K-j} \big), \tag{5}$$

where $H_n = \sum_{i=1}^{n} 1/i$ is known as the *n'th harmonic number* and we assume $H_0 := 0$. Putting everything together,

$$p(b_j) = p\big(\, n \in E(b_j)\,\big) \cdot p\big(\, b_j \mid n \in E(b_j)\,\big) \tag{6}$$

$$= \frac{1}{K} \big(\, H_K - H_{K-j}\,\big). \tag{7}$$

This implies a very strong bias towards using the highest base, which is further discussed in chapter **??**.

As a final check, does eq. 7 really defines a distribution — is it normalised? The distribution in eq. 7 is normalised if and only if

$$\sum_{j=1}^{K} \frac{1}{K} (H_K - H_{K-j}) = 1 \tag{8}$$

Multiplying both sides by K and reordering, we see that this is equivalent to
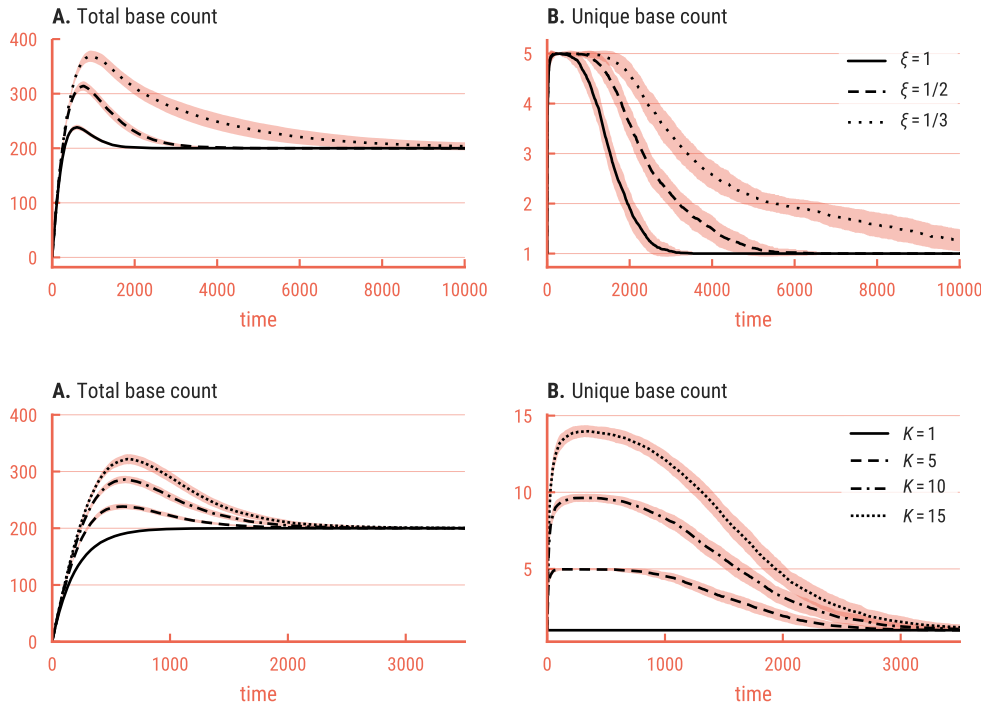
$$\sum_{j=1}^{K-1} H_{K-j} = K \cdot (H_K - 1), \tag{9}$$

where we used the convention $H_{K-K} = H_0 := 0$ to sum only up to $K - 1$. Equation 9 is in fact true: it is one of the basic recurrence relations on the harmonic numbers. We can therefore conclude that the distribution in eq. 7 is really a distribution.

As a final note, I have note been able to find any reference of this distribution, but I haven't searched extensively either. But if the reader happens to recognise this distribution — I would be very interested to learn more about it.

Properties of the additive base game

EFFECT OF ξ Recall that the parameter ξ determined which bases were favoured; base b was *favoured* if

$$s(b) > 0 \qquad \text{and} \qquad s(b) \geq \xi \cdot \max_b s(b), \tag{10}$$

A. Total base count

B. Unique base count

A. Total base count

B. Unique base count

for $1 \geq \xi > 0$. The additive base game was simulated with $\xi \in \{1, 1/2, 1/3\}$ and the results are reported in figure 1. The main effect seems to be that lower values of ξ slow down convergence. This is not surprising: when $\xi = 1/3$, an agent only favours a single base if its frequency is 3 times as high as that of any other base. This is a much weaker preference for frequent bases than with $\xi = 1$, in which case agents always use the most frequent base.

EFFECTS OF THE NUMBER OF SIMPLE NUMERALS   How does the number of simple numerals influence the game? Figure 2 summarises some experiments of the additive base game with $B \in \{1, 10, 20, 30\}$. Although these simple experiments do not yield strong conclusions, the convergence time is not strongly influenced by B. That is, certainly not in a power law fashion, like population size. Rather, the convergence time does not seem to depend on K. One reason for this is that convergence time in the Dirichlet-categorical naming game was also not found to be strongly influenced by K.