A social network model for the development of a ‘Theory of Mind’

Michael S. Harré
School of Engineering and IT, The University of Sydney, Sydney, Australia
E-mail: michael.harre@sydney.edu.au

Abstract. A “Theory of Mind” is one of the most important skills we as humans have developed; it enables us to infer the mental states and intentions of others, build stable networks of relationships and it plays a central role in our psychological make-up and development. Findings published earlier this year have also shown that we as a species as well as each of us individually benefit from the enlargement of the underlying neuro-anatomical regions that support our social networks, mediated by our Theory of Mind that stabilises these networks. On the basis of such progress and that of earlier work, this paper draws together several different strands from psychology, behavioural economics and network theory in order to generate a novel theoretical representation of the development of our social-cognition and how subsequent larger social networks enables much of our cultural development but at the increased risk of mental disorders.

1. Introduction
Imagine if every new-born child had to learn from scratch how to tie a knot, ride a bicycle or even speak the local language without the help of their parents and teachers. What we know, how we learn and how we reason is not just a matter of sitting down and using our brains, however large they may be, to solve a problem, without the supportive network of relationships that guide our learning there are just too many different ways for us to go astray. But what is the relationship between our innate abilities, conferred by our genes, and the development of the social networks in which we live? At the level of the species, Robin Dunbar suggested that the average social group size an individual lived in is related to the size of the neocortex by the following relationship [1]:

$$\log_{10}(N) = 0.093 + 3.89\log_{10}(C_r)$$  (1)

where \(N\) is the mean group size across a species and \(C_r\) is the ratio of the neocortex size to the rest of the brain across the species. The ratio of neocortex to the rest of the brain was used to allow for those species that need comparatively large brains for the management of their (relatively) large bodies, e.g. whales have large brains in part because they have large bodies that require considerable brain power to co-ordinate. A consequence of this model was the prediction that the average human social group size should be somewhere between 100 and 230 (95% CI), commonly approximated at about 150. At the level of the individual though, rather than at the species level, Dunbar and his colleagues have recently shown that the size of the Orbital Pre-Frontal Cortex (oPFC) region of the human brain linearly correlates with the size of an individual’s social network [2].
2. Developing a Theory of Mind

ToM in infants is often measured in several different ways, but the example I will use, called the Sally-Anne test [3], is one of the first developed and is sufficient for the ideas I want to illustrate by example. First used in 1983, the Sally-Anne test uses two dolls that a child, typically between 3 and 9 years old, is introduced to: Sally and Anne. The child witnesses a scene that contains the two dolls, a basket and a box and then watches while Sally places a marble in the basket in clear view of Anne. Sally then goes for a walk (leaves the scene) and Anne removes the marble from the basket and places into the box. When Sally returns the child is then asked: Where will Sally look for the marble? The correct answer, i.e. in the basket and not in the box, requires the child to recognise that Sally has different states of knowledge to that of the child (or that Sally can has inaccurate knowledge of the world). Autistic children often have difficulty passing this test because they are unable to attribute mental states to Sally that are different from their own, this is often interpreted as a failure to properly develop a ToM.

This test requires a child to have some model of the internal workings of another person’s mind and that the model needs to accurately reflect that person’s state of mind rather than the state of mind of the child. Figure 1 represents four different levels of modelling that a child might have. These are the basic steps by which a child forms a ‘link’ between themselves, and another through an increasingly better understanding of the other person’s point of view.

![Figure 1. Increasing levels of complexity in the development of a ToM. A. $C_1$ sees the world as simple objects. B. $C_1$ attributes an incorrect state of knowledge to $C_2$. C. $C_1$ attributes a correct state of knowledge to $C_2$. D. Both $C_1$ and $C_2$ attribute states of knowledge to each other, including states of knowledge regarding each other.](image)

3. Network Theory for Small to Medium Societies

In this paper I use the Erdős-Rényi [4] network model where we have a set of $N$ nodes and $L$ links whereby a link forms between two nodes with probability $p$. The following relationships are well established for this model (ln is the natural log): $\ln(N) = l$ and $N^{-1} \ln(N) = p$ where $l$ is the minimum expected number of links a single node needs to make and $p$ is the minimum probability that a link forms in order for the network to be fully connected$^1$

We want to describe the minimum necessary conditions under which a link is formed between two individuals in a small to medium society. Every individual $i$ has a set of $m$ behavioural attributes denoted by a vector $b_i = [x_1, x_2, \ldots, x_m]$, $x_k \in \{0, 1\}$ containing binary behaviours (present-not present). The probability that two people $i$ and $j$ have matching behavioural vectors is a probability $p = q^m$ where $q$ is the probability that a single behavioural attribute is a match. The probability that a person is able to match their $b_i$ to that of $l$ others is $p^l = q^{ml}$.

$^1$ Throughout we assume the caveat: with high probability
In a network of size $N$, each person needs to have minimum number of links given by $l = \ln(N)$ in order for a group of individuals to become a fully connected community, so each individual needs to connect with a minimum probability $p = q^{\ln(N)m}$. But this minimum is also known to be given by $p = N^{-1} \ln(N)$, so we have: $N^{-1} \ln(N) = q^{\ln(N)m}$ and solving for $m$ gives:

$$m = \frac{\ln(\ln(N)N^{-1})}{\ln(q) \ln(N)}$$

(2)

Figure 2. Behavioural complexity $m$ as a function of minimum connected network size $N$ and probability $q$ of accepting any single behavioural attribute.

This equation relates the maximum size of the behavioural repertoire $m$ (the behavioural complexity of each individual) to the group size $N$ and the minimum probability $q$ that a single behavioural trait is ‘acceptable’ between two members of the social group. If these conditions are met the group becomes a community in which all members of the group are connected. A plot of this function is shown in Figure 2. Note that $\frac{-\ln(\ln(N)N^{-1})}{\ln(N)}$ is a smoothly varying function with a limit of 1 as $N \to \infty$. In the discussion that follows this function is essentially a scaling factor over the $N$ of interest with a nearly constant value in the range of $c \in [0.63, 0.76] \simeq 0.7$ for $N \in [4, 3 \times 10^3]$ and so we focus on $m = -c \ln(p)^{-1}$. This is expressed as an equality but it represents a threshold, for $m$ larger than $-c \ln(q)^{-1}$ a connected network is highly unlikely to form, for $m$ smaller than $-c \ln(q)^{-1}$ a connected network is highly likely to form so the region of interest is $m \leq -c \ln(q)^{-1}$. A connected network is a social community of individuals whereby all members have at least one connection with another individual, a community has no loners.

4. Discussion

We appear to have some of the largest and most diverse social networks of any species on earth but we also have some of the most complex behaviour. This work argues for a mentalising ability that mediates the relationship between behavioural complexity and social network size. This mentalising is expressed as the probability that we can ‘comprehend’ each of the behaviours of another and the size of the social networks that can then form.

This is achieved by considering another person’s state of knowledge and constraints and imagining how you might act in the same situation. Such mentalising reduces the probability of a misunderstanding and has been shown to be key in ToM and social network size [2]. The left hand side of Table 1 shows how the number of link formations ($\ln(N)$), which is dependent on an individuals ability to accurately mentalise regarding other people’s behaviours, dictates the maximum social group size ($N$).
\[
\begin{array}{|c|c|c|c|c|c|}
\hline
N & \ln(N) & \ln(N)N^{-1} & [N_1, N_2] & [d\ln(N_1), d\ln(N_2)] \\
\hline
7 & 1.94 & 0.278 & [3, 5] & 1, 1.46497 \\
20 & 3.00 & 0.150 & [9, 15] & 2, 2.46497 \\
55 & 4.01 & 0.073 & [27, 45] & 3, 3.46497 \\
150 & 5.01 & 0.033 & [81, 135] & 4, 4.46497 \\
400 & 5.99 & 0.015 & [243, 405] & 5, 5.46497 \\
1100 & 7.00 & 0.006 & [729, 1215] & 6, 6.46497 \\
1.30 \times 10^{52} & 120 & \approx 0 & [2187, 3645] & 7, 7.46497 \\
\hline
\end{array}
\]

Table 1. **Left:** For a fully connected network of size \( N \), the minimum expected links per node \( \ln(N) \) and the minimum probability of a link being formed \( N^{-1} \ln(N) \) are shown in approximate discrete steps of links. **Right:** A connectivity table scaled by a constant \( d = \ln(3)^{-1} \), for maximum group sizes in the range \([N_1, N_2]\) the scaled expected connectivity per person is in the range \([d\ln(N_1), d\ln(N_2)]\).

The right hand side of this table shows the scaled connectivity required for networks of maximum sizes between \( N_1 \) and \( N_2 \). This table reflects the work of Dunbar and colleagues in understanding the multi-layer nature of human social networks [5]. In this work multilayered social networks increase in size with constant scaling factor slightly larger than 3. A range of group sizes they discuss are approximately of size \((N)\): 3-5 (support clique), 9-15 (sympathy group), 30-45 (overnight camps), 150 (clan), 500 (megaband) and 1000-2000 (tribe).

Other numbers were given based on a qualitative analysis of the data: \( S_0 = 1 \) (the self), \( S_1 = 4.6 \), \( S_2 = 14.3 \), \( S_3 = 42.6 \), \( S_4 = 132.5 \), \( S_5 = 566.6 \) and \( S_6 = 1728 \) with an average scaling between groups of \( S_i/S_{i-1} \) = 3.52. Note that the first four groupings are a reasonable approximation with \( N_2 \) in the table above, offering a potential solution to a question Dunbar et. al. ask at the end of their paper, why discrete scales and a factor of 3? Because individuals that can, on average, be expected to maintain between \( d\ln(N) \) and \( d\ln(N) + 0.5 \) links and as \( d\ln(N) \) is in discrete steps it has a maximum fully connected social group size of \( N \) that increases stepwise.

Finally, Steven Pinker [6] has recently commented on an on-going debate regarding traditional evolutionary theory where genes act solely at the individual organism level and ‘group selection’ where genes act at the level of benefiting a group of individuals and not the individual itself. For example group selection argues that self-less altruism can be positively selected for because it benefits the group while disadvantaging the individual organism. The current paper suggests that evolutionary pressure comes to bear on our mentalising abilities, and that the resultant larger groups are probably competitively advantageous. If this means understanding that someone needs help, and that we would expect help were we in the same position, then this ‘perspective taking’ ability can be selected for at the individual level.

[1] R.I.M. Dunbar. Neocortex size as a constraint on group size in primates. Journal of Human Evolution, 22(6):469–493, 1992.
[2] J. Powell, P.A. Lewis, N. Roberts, M. García-Fiñana, and RIM Dunbar. Orbital prefrontal cortex volume predicts social network size: an imaging study of individual differences in humans. Proceedings of the Royal Society B: Biological Sciences, 2012.
[3] H. Wimmer and J. Perner. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception. Cognition, 13(1):103–128, 1983.
[4] P. Erdős and A. Rényi. On the evolution of random graphs. Akad. Kiadó, 1960.
[5] W.X. Zhou, D. Sornette, R.A. Hill, and R.I.M. Dunbar. Discrete hierarchical organization of social group sizes. Proceedings of the Royal Society B: Biological Sciences, 272(1561):439–444, 2005.
[6] S. Pinker. The false allure of group selection. Edge, June 18, 2012.