"Businessmen always meet twice in their lifetimes."

Dictum of Zürich merchants in the 15th century

When Will We Meet Again?

Regularities of Social Connectivity As Reflected in Memory and Decision Making

Thorsten Pachur, Lael J. Schooler, & Jeffrey R. Stevens

Pachur, T, Schooler, L.J., & Stevens, J.R. (2012). When we meet again: Regularities in social contact dynamics reflected in memory and decision making. In R. Hertwig, U. Hoffrage, & the ABC Research Group (Eds.), *Simple heuristics in a social world* (pp. 199-224). Oxford: Oxford University Press.

In 1967, the social psychologist Stanley Milgram reported a fascinating observation. For a study on social connectivity, he randomly picked several individuals in Wichita, Kansas, and asked them to get a message delivered to a particular target person in Cambridge, Massachusetts—a distance of more than 1,300 miles. There were two catches. First, Milgram provided no address but only the name and some basic information about the target person, such as the area where he lived and his profession (the target person was the same for everybody). Second, people could only forward the message to someone they knew on a first-name basis. That is, the task for each person in the chain was to pass on the message to someone she thought would be most likely to know the target person or know someone else who might know him. The astonishing result was not only that many of the messages reached the target, but also that the number of intermediaries needed turned out to be rather small, ranging between 2 and 10 (with a median of 5).[1]

Milgram's (1967) findings became known as the small-world phenomenon or six-degrees of separation because it takes six separate journeys to get from the sender to the target via five intermediaries. His results suggest that the social world possesses regularities that allow most people to reach any other person using a relatively small number of steps. Researchers have only recently begun to understand the nature of these regularities in greater detail. Watts and Strogatz (1998) identified conditions under which networks are both "small worlds", in which the average number of connections between all agents in the network is rather low, and, at the same time, display tight clustering consisting of densely connected subgroups. Specifically, some connections must link members of a cluster to a randomly selected member from the entire network. Numerous investigations have followed, exploring the nature of social environments.

Most studies of social networks have concentrated either on what we term "network structure"—the distribution of connections across members of a network (e.g., Watts, 2003)—or on the size and hierarchical structure of social environments (e.g., Dunbar, 1992, 1998; Hill & Dunbar, 2003; Zhou, Sornette, Hill, & Dunbar, 2005). In our work, we will focus on less well-studied features of networks: how *connection strengths*—that is, how often one has contact with a particular person—are distributed across network members as well as the *network contact dynamics* to which the distributions give rise. Although it is usually not mentioned in discussions of the small-world phenomenon, Milgram (1967; see also Travers & Milgram, 1969) had already observed a particular structure in connection strength. Specifically, from the target person's network members only very few served as the final step for transmitting the messages. In fact, from the 64 messages that the target person received, 31 were transmitted through the same *three* network members—a very small number given that network analyses usually estimate the size of a person's network to be 150 or higher (e.g., Killworth, Bernard, & McCarthy, 1984). In other words, the target person's distribution of connection strengths was highly skewed. Here we examine more generally to what extent a skewed distribution characterizes social contacts. In addition, we explore how the probability of contact with another person from one's social network can be predicted. In particular, we examine the relationship between aspects of past contact and the probability of future contact and how this relationship can be described mathematically. Our investigation is motivated by previous analyses suggesting that characteristic patterns of memory performance are paralleled by patterns in the environment (Anderson & Schooler, 1991). These analyses

[1] Milgram's original empirical evidence can be criticized on various grounds. For instance, it could be objected that this figure is based on only the completed transmissions. However, an analysis considering all initiated transmissions yielded an only slightly higher estimate (Travers & Milgram, 1969). A number of further problematic aspects in Milgram's original data were highlighted by Kleinfeld (2002). All these objections notwithstanding, Milgram's finding of very short average path lengths was recently replicated in a large international study looking at email communication (Dodds, Muhamed, & Watts, 2003).

highlighted the role of frequency, recency, and the spacing of events for predicting both future events and memory performance.

Do such patterns exist for social environments as well? The design underlying some communication systems seems to be based on the assumption that they do. In many types of email software, email addresses from one's contacts are ordered by frequency of past contacts, assuming that most frequent previous contacts are most likely to be contacted in the future. Similarly, mobile phones allow for a quick access to the numbers of the most recent calls, implicitly assuming that these are also most likely to be used in the future.

If statistical regularities between past and future contact indeed exist and if these regularities correspond to those found for memory retrieval, they might be important variables for studying and understanding aspects of social decision making, such as how well simple heuristics perform in a social world and how they might exploit memory retrieval. In the following, we report the results of an empirical study in which we found evidence for strong statistical regularities in human social contact. We then discuss the potential implications of the regularities for social decision making and how they might be exploited by simple heuristics, especially those associated with cooperation. Finally, we explore, based on a formal model of memory, to what extent the familiarity signal generated by traces of previous contacts can be used to reliably predict future social contact.

How the Past Predicts the Future: Three Regularities

Based on a mathematical model for patterns of library borrowing, Anderson and Milson (1989) developed a memory model that assumes that the availability of memory traces matches patterns in the traces' previous retrieval needs. The key hypothesis underlying the model is that certain characteristic memory phenomena represent efficient responses to information that the cognitive system typically needs to retrieve. More specifically, the phenomenon that items are better recalled if they occurred more frequently or more recently reflect typical occurrence patterns in the environment. Moreover, if these occurrences were spaced over time relatively equally (rather than massed/clumped) one typically observes better recall at longer delays, but the reverse at shorter delays, where massed occurrences tend to produce better recall. In a later set of ecological analyses, Anderson and Schooler (1991) provided support for this hypothesis. They demonstrated regularities between the probability of a certain word's occurrence and the pattern of previous uses of the word in the headlines of the *New York Times*.

In addition—and crucial for the focus of this chapter—the results of another study in Anderson and Schooler (1991) hinted that such systematic relations might even exist in social contacts. Past social contact strongly predicted future contact in very specific ways. Analyzing three years' worth of email received by John Anderson (JA), three regularities emerged relating to the frequency, recency, and spacing of contact.

Frequency Effects

Anderson and Schooler (1991) first investigated how the frequency of past contacts influenced future contact. They found that the probability p of JA being contacted by someone on day t increased linearly with the number of days f on which he had been contacted by this person in the previous w days: $p = b_0 + b_1 f$. In this equation, the intercept b_0 represents your baseline social activity (i.e., the probability that a contact occurs with someone from your social network with whom there was no contact in the past w days); and b_1 reflects the degree to which the probability increases with more frequent previous contacts. Thus, the probability of future contact increases proportionally with the frequency of previous contacts.

Recency Effects

Next, Anderson and Schooler (1991) tested how the time since last contact (recency) influenced future contact. Like frequency, recency strongly predicted future contact. Specifically, there was a power relation between the odds $o = p/(1-p)$ that JA was contacted by someone and the number of days r since his last contact with that person (on day t-r): $o = b_0 r^{-\alpha}$. Again, the intercept b_0 in this equation represents baseline social activity. The parameter α reflects how quickly the probability of contact decreases as the number of days since the last contact increases. The power relation suggests that the odds of future contact decreases considerably with the time since last contact. If the last contact occurred very recently (e.g., yesterday), you are very likely to have contact with this person again soon. With less recent contact (e.g., several days ago), the probability of contact decreases dramatically. The power function also implies a scale-free relationship; that is, the probability of contact shows a similar pattern if recency is scaled in minutes, days, or years.

Spacing Effects

Finally, Anderson and Schooler (1991) examined how the spacing of contacts might predict future contact. Take two persons with whom you had the same number of contacts previously but for one person the contacts were spread out over time, whereas for the other the contacts were clumped together. The probability of contact soon after the last contact is higher for the latter, whereas the reverse is true for the probability of contact at longer lags after the last contact. This pattern can be visualized as follows. Assume that p is plotted as a function of recency of last contact (holding f constant), separately for *massed* (*or clumped*) contact (say, occurring on consecutive days) and *spaced* contacts (occurring days apart). At short recencies, p was higher for massed than for spaced contacts, but at longer recencies the pattern reversed (i.e., the lines crossed). To illustrate, assume you have met someone four times during the last month and the last encounter was only a few days ago. The chances are higher that you will see that person again if the four encounters occurred on consecutive days than if they occurred only once per week. If the last encounter was less recent, by contrast, the reverse is true: you will likely see that person again if the four encounters occurred only every week rather than if they occurred on consecutive days.

Taken together, these regularities suggest that the probability of having future contact with a person via email is highly predictable from how frequently and recently previous contacts occurred and how these contacts were distributed over time, paralleling results found for patterns of word usage. Crucially, the same systematic relationships to frequency, recency, and spacing also exist for memory performance (e.g., Cepeda, Vul, Rohrer, Wixted, & Pashler, in press; Glenberg, 1976; for an overview see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). Specifically, memory performance for an item increases proportionally to the frequency with which the item has been studied; it decreases as a power function of how long it has been since the last study; and it is also affected in a characteristic way by the spacing of the study trials. Items learned through massed study (i.e., occurring in quick succession) tend to be more easily retrieved than items learned through spaced study (i.e., spread out over time) at short lags after the final study trial. In contrast, at longer lags, the picture reverses, such that items learned through spaced study tend to be better recalled than those learned through massed study. Thus, the performance function for items learned through massed study decays more quickly than the one for items learned through spaced study.

Anderson and Schooler's (1991) results indicate a correspondence between patterns in memory retrieval and patterns in which social information (word usage and email contacts) occur in the environment. Memory might thus be adaptive in the sense that it matches the availability of memory content to the needs posed by the environment. Given the frequency and importance of social information, the social environment may be a powerful force in shaping memory systems.

Although Anderson and Schooler's (1991) analysis of email contacts hints at regularities in social environments and, although these regularities are similar to those in other domains, it is unclear to what extent these results generalize. Are the regularities observed in email contacts representative of a broader structure of social communication that also includes face-to-face interactions, phone calls and other communication modes? After all, even for heavy users, email communication makes up only about 25% of people's social contacts (Copher, Kanfer, & Walker, 2002). The issues of whether regularities exist in social environments, how they can be described, and whether they match regularities in other domains (and, by implication, characteristics of memory performance) are important for a number of reasons. First, systematic patterns in the actual distribution of social encounters should have implications for analyses of strategies for social interactions, which often assume that contact between agents is equally probable. Second, to the extent that simple heuristics in the social world exploit a strong correspondence between patterns in memory performance and social encounters, studying the structures of social environments is informative for the study of simple heuristics.

Do the Regularities Also Exist in Social Contacts?

To examine the statistical structure of social contacts, we analyzed records of social contacts across a period of time. In contrast to Anderson and Schooler (1991), who considered only daily email, we also recorded contacts via face-to-face interactions, phone calls, and letters. The recorded contacts included being contacted (e.g., receiving a phone call), contacting others (e.g., giving someone a call), and unplanned encounters. The study involved the records of the daily contacts by one of us (TP) over 739 days. All results reported for this study were replicated in a study involving 40 participants who recorded their social contacts over a period of 100 days (Pachur, Schooler, & Stevens, 2010). Given the relatively long study period, our goal was to make record keeping manageable by striking a balance between the reliability and the level of detail of the records. Therefore, social contact was defined as all face-to-face or phone conversations lasting at least five minutes and all communication conveyed electronically or on paper of at least 100 words in length—thus only rather "intense" contacts were recorded. In addition, we tracked only whether a contact with a particular person occurred on a particular day, but not the length and the medium of contact, nor the direction of contact (i.e., which side initiated it), or how many contacts with a person occurred on each day (but see Pachur et al., 2010, who found that the results hold across different communication channels—such as face-to-face, phone, and email—and independently of the direction of contact).

Contacts were recorded in an electronic diary. Over the period of 739 days, contacts with 351 different persons were recorded, with an average of 7.04 contacts per day (*SD* = 3.72), ranging between 1 and 23. Figure 1a shows that the distribution of probability of contact (defined as the proportion of days in which a contact occurred) across network members was highly skewed: Although for some network members contact probability was very high, such network members were extremely rare. For most members of TP's network, the probability of contact during the study period was smaller than 5% on any given day. The median probability of contact was 0.4% and the probability was 20% or higher for only 1.1% of TP's network members (4 out of 351). Figure 1b plots the data as a cumulative distribution function—that is, it shows for the different levels of contact probability p, the proportion of network members with contact probability p or larger. The distribution is very well captured by a power function (R^2 = 0.99), which, in the log-log plot, yields a straight line.[2] The estimated value for power exponent α in the noncumulative distribution was 1.62. Interestingly, our finding of a power-law distribution for the probability of contact (i.e., the strength of connections in a network) mirrors what has been found for the distributions of the number of connections in many natural networks, such as semantic, computer, or protein networks. For instance, Barabási and Albert (1999) observed that the connections are usually not distributed uniformly among all nodes. Instead, only few nodes have a very large number of connections

[2] Although Pachur et al. (2010) also observed highly skewed distributions of contact probability in their study of 40 participants' social contacts, the majority of their participants were best described by an exponential function. An exponential function fitted TP's data worse than a power function (R^2 = 0.97).

to other nodes (so-called "hubs'), whereas the large majority of nodes have only a moderate number of connections.

Clearly, our data suggest that frequent contact occurs with only a very small number of network members, yet with most members contact is relatively rare. In addition, contacts often occur in clusters across time. For illustration, Figure 1c shows the contact patterns of the 20 network members with the most contacts across the period of 739 days. For some network members the contacts are spread out relatively equally, whereas for others they occur in clusters. That is, several days without contact are interspersed with days with very frequent contacts. As we will see shortly, differences in how contacts are distributed can influence contact probability.

Can we predict future contact based on aspects of past contact? Specifically, are there similar regularities in contact strength in terms of how the frequency, recency, and spacing of past contact influence the probability of future contacts? To answer these questions, we conducted the following analyses. For the effect of frequency, we determined those network members with whom, in a time window of w days, f contacts occurred and calculated the proportion for which a contact occurred on day $w + 1$. For the effect of recency, we determined those network members with whom, in a time window of w days, the last contact occurred r days ago and calculated the proportion for which a contact occurred on day $w + 1$. For all analyses, we used a time window of $w = 30$ days.

Frequency Effects

In what way is the probability of interacting with a person related to the number of previous contacts with the person? We calculated the probability that a contact occurred to a person on the 31st day as a function of the number of contacts with that person in the preceding 30 days. Figure 2a shows the probability of contact depending on the number of previous contacts. We find that probability of contact increases linearly with the frequency of previous contacts, $p = -0.01 + 0.03f$ ($R^2 = 0.94$). In other words, the pattern found by Anderson and Schooler (1991) for email contact generalizes extremely well to a broader range of interactions: the frequency of past contact predicts the probability of future contact.

Recency Effects

As a further step, we asked how the probability of having contact with a network member relates to the recency of the last contact with the person. Figure 2b shows the odds of contact, o (= [$p/(1-p)$]), as a function of the recency of the last contact, with both variables plotted on a log scale.[3] We again find a highly systematic relationship, with a linear fit implying a power relationship between the untransformed data. In fact, the untransformed data are well described by the power function $o = 0.46\, r^{-0.97}$ ($R^2 = 0.91$). Figure 2b, however, also reveals some systematic deviations from the power function, with the probability of contact spiking at r = 7, 14, 21, and 28 (indicated by arrows). The location of these spikes suggests that they reflect weekly cycles in human activity.

Combined Frequency and Recency Effects

The last contact with a person we have encountered often will have occurred, on average, more recently than the last contact with a person we contacted only rarely; frequency and recency (i.e., the number of days since last contact) of contact are thus correlated. In our data, this correlation was $r = -0.25$—that is, the higher the number of past contacts, the lower the number of days since last contact.[4] To disentangle the influence of recency and frequency we analyzed the relation between the probability of contact and recency separately for

[3] Note that Anderson and Schooler (1991) proposed that this "retention curve" is well described by a negatively accelerated power function, $Y = X^{-\alpha}$. As power functions are unbounded above, Y cannot be a bounded measure, such as probability of contact. Therefore, we used odds instead, which are unbounded above like a power function.

[4] To compute this correlation, we determined for each 5,052 times that contact with someone occurred, the number of days since the last contact as well as the number of contacts to that person in the time window.

different frequencies of contact. Figure 2c shows the probability of contact as a function of recency for network members with high frequency of contact (defined as 6-15 contacts in the previous 30 days) and those with low frequency of contact (1-5 contacts). We find similar patterns irrespective of the number of past contacts (although, of course, the contact probability was higher for high frequency contacts). In other words, for different contact frequencies, the decay parameter of the corresponding recency curves is similar. The best fitting functions were $o = 0.26r^{-0.77}$ ($R^2 = 0.96$), and $o = 2.22r^{-1.19}$ ($R^2 = 0.87$) for low and high frequencies, respectively.

Spacing Effects

As a last step, we turn to a more subtle predictor of social interaction: the spacing of past contacts. For the purpose of this analysis, we selected cases with exactly two contacts ($f = 2$) in the past 30 days and considered the probability of contact on the 31st day. We distinguished two spacing patterns: *massed contacts*, defined as cases in which the two contacts occurred on two consecutive days (short lag); and *spaced contacts*, defined as cases with a lag between the two contacts of between 2 and 28 days (long lag). In Figure 2d, the probability of contact is plotted as a function of recency, separately for massed and spaced contacts (results were robust when higher values of f were used). The probability of contact was analyzed as a function of recency of last contact using running averages and binned according to the following scheme. The data points for days 3 through 28 are running averages over five-day bins, including data from the two preceding and the two subsequent days (e.g., the running average for day 3 was computed as the average over the days 1-5). Days 2 and 29 are running averages over three-day bins, and days 1 and 30 are moving windows over one-day bins.

How would the probability of contact vary as a function of the recency of contact for massed and spaced contacts, respectively? Figure 2d shows that, as expected, there was an interaction between recency of last contact and how past contacts were distributed over time, indicated by the lines for massed and spaced contacts crossing. At short recencies, the probability of contact is higher for massed than for spaced contacts. At longer recencies, however, the pattern reverses. Between 5 and 23 days since last contact, the probability of contact tends to be higher for network members with spaced contact than for those with massed contact. Note that we thus find that social contact displays regularities that mirror spacing effects found for memory: at short lags, memory is better for items learned through massed rather than spaced study; at long lags the reverse is true. This parallel might be taken to indicate that our memories encode the statistics of the world—social and nonsocial alike—to make information available according to the dictates of these statistics.

In sum, our analyses show that the probability of future contact (a) is highly skewed across the people one knows—contact is likely with only very few network members, and (b) follows systematically from the frequency, recency, and spacing of past social contacts. All of these findings were replicated in the study involving 40 participants (Pachur et al., 2010). The statistical structure of social contacts thus shares important characteristics with patterns of word usage in language and with patterns of email contacts (Anderson & Schooler, 1991). This suggests that efficient information retrieval systems operating in these domains may require similar designs. For instance, mobile phones and other communication systems could gain efficiency by making phone numbers and email addresses available as a function of not only frequency and recency but the spacing of past contacts as well.

What Mechanisms Might Generate the Structures in Social Environments?

The observation that social environments display strong regularities begs the question of which mechanisms might generate them. For instance, take the highly skewed distribution in the probabilities of contact, where frequent contacts occur only with a very small group of network members (Figure 1a). Given that this distribution displays some characteristics of a power-law distribution (Figure 1b) one might consider the *Yule process* as one candidate (cf. Newman, 2005; Simon, 1955). The Yule process has been proposed as an explanation for power laws in domains such as city sizes (Simon, 1955), paper citations (Price, 1976), and links to pages on the World Wide Web (Barabási & Albert, 1999). The basic idea, often termed a "the-rich-get-richer process" or also "Matthew effect', is that if an entity has more of some property (e.g., money, citations, population), it will tend to gain proportionally more of this property as compared to an entity with less. The Yule process, which yields a power-law distribution, is certainly reasonable in the context of social contact. Applied in the context of social contacts, a Yule process would yield a power-law distribution when three conditions are met. First, new members are, on occasion, added to your social network. Second, in between the addition of one new network member and the next, you have further contacts with your network members and for each, there is some probability that she will be the next person with whom you have contact. Third, the probability that you will have contact with a network member is proportional to the frequency of past contact. Figure 2a, which shows that for our study the probability of contact is proportional to frequency of past contact, appears to support this last condition.

But not only is this proportionality condition empirically supported, it also makes intuitive sense. Each time you see someone there is some probability that you will arrange a future meeting. Each time someone sends you an email there is some probability that you will correspond again. Simon (1955) suggested that the generality of these conditions explains why power-law distributions are so common. The Yule process also underlies Barabási and Albert's (1999) preferential growth model of scale-free network structures:[5] hubs exist because new contacts attach preferentially to nodes having many contacts already.

Mind and Environment: What Causes What?

The observed correspondence between the structure of the environment and characteristics in memory performance begs the question of causality: are characteristics of memory due to the structure of the environment, or is the environmental structure due to characteristics of memory? The evolutionary biologist Lewontin (1983; see also Brandon, 1996; Laland, Odling-Smee, & Feldman, 2000) distinguished two ways of understanding the relation of organisms to their environments: an adaptationist view and a contructivist view. In the *adaptationist view*, the environment molds species by posing problems for them to solve. Anderson and Schooler's (1991) analysis of the *New York Times* headlines exemplifies an adaptationist analysis, as a headline reader has no influence over what appears there and thus passively experiences the environment (Schooler, Serio-Silva, & Rhine, 2010). In the *constructionist view*, by contrast, organisms both are shaped by their environments and also actively shape them. As Lewontin puts it, organisms "create a statistical pattern of environment different from the pattern in the external world. Organisms, by their life activities, can damp oscillations, for example in food supply by storage, or temperature by changing their orientation or moving" (Lewontin, p. 281). In an analysis of verbal interactions of preschool children (Anderson & Schooler, 1991), what a child said in one utterance influenced what she would hear in the future. We argue that in our analysis of social contacts a constructivist view is more appropriate than an adaptationist view. TP's memory—through his actions—influences whom he contacts and the contacts he receives, in turn, influence which memory traces are made available. A striking aspect of the findings for constructed environments is that even though the person is actively shaping them (rather than simply responding passively to them), the same robust statistical regularities emerge as for the passively experienced *New York Times*. Thus, for the purposes here, we need not have a strong model of the direction of causality, because similar patterns emerge with and without feedback from the agent to the environment. Similarly, Pachur et al. (2010) observed the same regularities in contact probability irrespective of which side initiated the contact.

[5]In scale-free networks, the connectivity distribution—that is, the number of nodes that have a particular number of connections to other nodes—follows a power law.

Social Structure and Simple Heuristics

If, as our results suggest, the social world displays robust statistical structures, these structures may influence how well simple heuristics perform in this world and eventually also which heuristics developed in the first place. Moreover, given the correspondence between systematic patterns in the social environment and those in memory performance, regularities in social contacts should be reflected in the ease with which information about these contacts are retrieved from memory. As a consequence, simple heuristics might exploit this statistical structure via memory. In this section, we consider these possibilities. We briefly describe one example of a simple heuristic for cooperation and then consider how the structure of social environments might affect the performance of this and other heuristics. We first give an overview of research focused on the influence of network structure (e.g., small-world characteristics, scale-free networks) on the performance of simple heuristics. We then elaborate on how regularities in the contact strength and contact dynamics might influence the performance of simple heuristics. In addition, we discuss the implications for how social distance—which is related to social connectedness—influences willingness to cooperate.

Evolution of Cooperation

Imagine two thieves are caught by police after a heist. They are immediately separated from each other, and each is given the following deal: testify against your partner to reduce your sentence. If the other guy testifies against you, then choosing to testify results in fewer years behind bars. If the other guy does not testify, choosing to testify once again reduces the time in prison. No matter what the other guy does, individuals have an incentive to cheat the opponent, even though their combined time in prison would be minimized by both choosing not to rat each other out. This game is called the prisoner's dilemma (dubbed the "*E. coli* of social psychology", Axelrod, 1984, p. 28) and models cooperation when a sizable temptation to defect (i.e., not cooperate) exists (chapter 12; Rapoport & Chammah, 1965). In the prisoner's dilemma, one-sided cooperation (in this example, not testifying against your partner) always results in lower benefits for an individual than defecting (i.e., testifying), but mutual cooperation yields the largest total group benefit (Figure 3). If defecting is always beneficial for an individual, how could we ever expect cooperation to evolve?

A number of general mechanisms have been proposed that can lead to cooperation in the prisoner's dilemma. Though some of them, such as kin selection, apply to one-shot situations, most mechanisms apply to the repeated game (iterated prisoner's dilemma). Trivers (1971) proposed reciprocal altruism (here referred to simply as reciprocity) as one key account for the evolution of cooperation in repeated interactions. In reciprocity, cooperating is contingent on previous cooperation by your partner. A simple heuristic that implements reciprocity is tit-for-tat (TFT): start by cooperating, and then copy your partner's previous choice (Axelrod & Hamilton, 1981). Using TFT only requires memory of the partner's most recent response (but see Stevens & Hauser, 2004).

Crucially for our present purpose, a critical variable for TFT to evolve is the probability of encountering someone again in the future, P (Axelrod and Hamilton used the variable w, poetically dubbed "the shadow of the future"). Specifically, TFT can outperform pure defection as long as $P > c/b$, where c represents the cost of cooperating and b represents the benefit from receiving the altruistic act (Axelrod & Hamilton, 1981; Nowak, 2006). Thus, although TFT itself focuses on past behavior, the conditions under which the strategy can evolve depend on the probability of future interaction.

The $P > c/b$ rule for TFT is derived using the standard evolutionary game theoretic analysis of cooperation. This analysis, however, relies on rather strict and unrealistic assumptions about interaction patterns between players. In particular, it assumes a well-mixed population in which every player has an equal chance to play against (i.e., interact with) every other player (Maynard Smith, 1982). Under these conditions, small group sizes and many interactions are required for TFT to be successful. As research on the small world phenomenon and our contact study have revealed, however, interaction patterns in real networks often display regularities that contradict these assumed conditions. For instance, we found that the distribution of contact probability is not uniform across players. This insight has important implications for understanding the performance of heuristics—and heuristics for cooperation are no exception. In the following, we first review how network structure influences cooperation and then consider the role of contact probability.

Cooperation and network structure. The assumption in evolutionary game theory that every player has an equal chance to play against every other player emphasizes only pair-wise interactions. Social network analysis, by contrast, embeds these pair-wise interactions in a larger matrix of interactions, with possible synergistic effects.[6] How is the evolution of cooperation influenced by the structure of networks, which vary in how individual network members are connected? In Table 1 we consider three types of network structures: regular, small-world, and scale-free networks (Newman, 2003; Watts, 2004). In regular networks, each agent interacts only with a fixed set of neighbors, typically nearby in space. When agents interact in local clusters with the same number of other partners, stable cooperation can exist using the simplest heuristic: always cooperate (Nowak & May, 1992; Ohtsuki, Hauert, Lieberman, & Nowak, 2006). For this unconditional cooperation, no memory or recognition of the partner's strategy is needed (although regular spatial structure stabilizes TFT as well; Pollock, 1989).

Although unconditional cooperation and TFT can evolve in regular networks, they do not represent a very realistic social structure. No one interacts solely with his or her neighbors. Watts and Strogatz (1998) added an interesting twist to the assumption of a fixed number of contacts in regular networks. By "rewiring" a few of the connections—that is, breaking the connection with one network member and randomly assigning it to another—shortcuts across the network are generated. In this structure, members interact primarily with their neighbors but sometimes have connections with more distant members (for a similar treatment of imposing shocks on neighborhood structures, see chapter 15). As mentioned above, Watts and Strogatz showed that this structure gives rise to the small-world characteristics that Milgram (1967) observed in his classic study; namely that the average number of connections between agents in a network is rather small. The high level of clustering in regular graphs allows cooperation to be maintained, and cooperation is even maintained when a small number of random connections are added. Increasing the number of random connections between members, however, degrades the evolution of cooperation (Masuda & Aihara, 2003; Watts, 1999). The likely reason is that the introduction of defectors through these random connections inhibits cooperation from gaining a strong foothold in small worlds.

Although small-world networks provide a more realistic pattern of interaction for social networks than regular networks, Barabási and Albert (1999) found evidence for an interesting subclass of small-world networks: scale-free networks. In the original formulation of small-world networks, all agents have the same number of connections (like in a regular graph). Scale-free networks, by contrast, have a power-law distribution of a number of connections—that is, some members have a very large number of connections ("hubs"), whereas others have only a few. Although not all small-world networks have scale-free characteristics (Amaral, Scala, Barthélémy, & Stanley, 2000), some do (e.g., sexual contact networks; Liljeros, Edling, Amaral, Stanley, & Aberg, 2001).

Importantly, Santos, Rodrigues, and Pacheco (2006) showed that the evolution of cooperation is influenced by whether a network displays scale-free characteristics or not.

[6] Social network analysis allows for the evaluation of the influence of connections beyond the target member. For instance, one can explore not only individual A's interactions with B and C but also B and C's interactions with each other. Depending on the networks structure, B and C may or may not interact, which can influence the payoffs of cooperation.

Specifically, unconditional cooperation was more stable in scale-free networks than in regular networks, even with large temptations to cheat. Once hubs were occupied by cooperators, defectors had difficulty spreading. In addition, Santos et al. showed that cooperation was greatly facilitated by network growth (i.e., the size of the network increased at each time step) and preferential attachment (because it allows cooperators that are hubs to make more connections and thus attract more cooperators). This probably also applies to the evolution of TFT, but we know of no studies examining this possibility.

Cooperation and probability of contact. Compared to the influence of network structure on cooperation, the influence of the distribution of contact probability (i.e., the connection strengths within a network) has so far received much less attention. Most analyses assume an equal probability of contact across all connections, and this also holds for most analyses of the iterated prisoner's dilemma (Axelrod, 1984; Maynard Smith, 1982). One of the key findings of these analyses was the identification of the algebraic conditions for sustaining TFT (i.e., $P > c/b$).

Our analysis of probability of contact, however, suggests that the assumption that all agents face the same probability of interacting with each of the partners in their network is not very realistic. Instead, we found that individuals interact more with some network members than others. This has important implications for the evolution of cooperation in real social networks. Specifically, there will not be a single probability of future encounter, P, but an individual probability for each partner P_i (where i indexes each partner or class of partners). Our data allow us to calculate realistic values for the average P as well as P's distribution (e.g., Figure 1a) and to explore how TFT performs under these more realistic conditions. In our analysis, the average P was about 0.02. With such low probability of contact the benefits of cooperation need to exceed the costs by a ratio of at least 50:1.[7] As this benefit-to-cost ratio may not exist in many real interactions, a skewed distribution of contact probability limits the evolution of TFT and, by extension, the evolution of the whole class of strategies that depend on repeated interactions to succeed.

But if TFT is unlikely to lead to cooperation with low probabilities of contact, how to account for the emergence of cooperation? There are two possibilities. The first is to leave the class of strategies that require a minimum number of interactions and investigate the emergence of cooperation with strategies like Win-Stay, Lose-Shift (or Pavlov; Kraines & Kraines, 1989; Nowak & Sigmund, 1993) under a realistic distribution of contact probability. Alternatively, we can stay in the TFT family but build in sensitivity to differences in the probability of contact. For instance, rather than always employing TFT, individuals may cooperate with a particular partner based on the probability of future interaction: cooperate if $P_i > c/b$. In this version, cooperation may emerge for the small subclass of frequently encountered partners, but not for others. Such an approach, however, raises the interesting question of what kinds of cues agents can use to predict the probability of future contact. Below, we explore how memory might serve as one such cue.

Social discounting. With the predictions from reciprocity in hand, we might expect that probability of future encounter will be a good predictor of cooperative behavior. That is, individuals may use a strategy such as "help those with whom I interact frequently". Under the assumption that social distance (defined as the perceived distance between individuals in other social groups) is highly associated with probability of contact, Jones and Rachlin's (2006) results confirm this prediction. These authors measured cooperativeness by how much money people would forgo to share with another person as a function of the social distance to that person. As it turned out, cooperativeness decreased with social distance as a hyperbolic

[7] If $P > c/b$ and $P = 0.02$, then $0.02 > c/b$, that is, the ratio of benefits to costs needs to exceed 50:1. When the median or mode probability (median = 0.004; mode = 0.001) are used, the critical threshold for the benefit-to-cost ratio increases to 250:1 and 1000:1, respectively.

function (which is a special case of a power-law function). That is, cooperativeness dropped off steeply at short social distances, but the rate of decrease slowed at larger social distances.

Interestingly, the highly skewed distribution of probability of contact that we observed (see Figure 1a) might offer an explanation of the particular form of the relationship between cooperativeness and social distance (assuming that social distance and probability of contact are correlated). To appreciate the argument, it is important to note that Jones and Rachlin (2006) defined the social distance of a person as the *rank* of that person among "the 100 people closest to you" (p. 284). The hyperbolic discounting function could result from the fact that the social distance between, say the closest and the 5th closest person is much larger than between the 20th and the 25th closest person. As a consequence, social discounting might not follow a hyperbolic, but a linear function when social distance is not measured by rank but by probability of contact.

In summary, as for the role of network structure, cooperation can emerge in networks in which cooperators occur in clusters. Notably, cooperation occurs in regular networks in which the same individuals interact repeatedly, as well as in scale-free networks in which hubs are occupied by cooperators. A large number of random connections (as in standard small worlds), by contrast, seems to inhibit cooperation because defectors can invade the clusters. As for the role of probability of contact our data suggest, first, that the assumptions underlying existing models of cooperation do not reflect realistic structures in interaction patterns. In other words, evolutionary models of cooperation need to be reformulated based on actual interaction patterns. Second, our results show that the probability of future contact is quite low for most partners, which has particular implications for the feasibility of TFT. In fact, the standard version of TFT likely will not fare well in an environment with highly skewed distributions of contacts (and this probably applies to the entire class of strategies that require a minimum number of interactions with partners). Cooperation might still emerge with a version of TFT that is sensitive to the probabilities of future contact with different partners. Strategies such as "always cooperate with partners with whom I interact frequently" (e.g., friends, relatives, co-workers) may perform well in real social networks because they exploit important aspects of the interactions such as the frequency, recency, and spacing of contacts. This intuition is confirmed in an analysis of cooperativeness with individuals varying in social distance—socially "close" individuals (presumably those with high probability of contact) receive more cooperative benefits. In other words, probability of future contact might be a key variable in an individual's decision of whether to cooperate with another person or not. But how could a boundedly rational agent judge the probability of future contact? In the next section, we discuss how a memory system that reflects the statistical patterns in social contact can be exploited by simple heuristics.

Memory-Based Heuristics for Cooperation

Evolutionary models of cooperation typically ignore the cognitive building blocks that are required to implement strategies for cooperation (Stevens & Hauser, 2004). These building blocks include individual recognition, quantification, delayed gratification, and inhibitory control, to name a few. For our purposes, the temporal delay between interactions requires an ability to recall the previous behavior of partners. How might memory affect cooperation?

The influence of memory on the performance of simple heuristics is illustrated, for instance, in a study by Aktipis (2006). Using evolutionary simulations, she considered various heuristics for deciding whether to cooperate with another agent and tested them in terms of their ability to invade (i.e., evolve to be the majority) a population of defectors. The heuristics she tested were simply based on recognition of whether an agent had been encountered before or not. Aktipis distinguished between heuristics that only stored agents who defected at previous encounters and those that only stored agents that cooperated. In addition, she manipulated memory size by varying the number of previously encountered agents stored (cf. chapter 15). One key finding was that a heuristic that stores only cooperators (but not defectors) could invade a population of defectors, but only when memory was extremely small. Specifically, the heuristic was successful if only the first cooperator (i.e., a memory size of one) was stored. Importantly, note that this heuristic is highly conservative: it leads to cooperation with only one single agent—namely the one stored in memory. In light of our result that contact probability is highly skewed and frequent contacts occur with only a very small group of network members, such a highly selective heuristic may approximate, to some degree, the structure of actual social environments. In other words, even if an individual is embedded in a large social network, cooperation with only a very small number of network members may be appropriate if the environment is structured such that most (and frequently repeated) interactions occur with only very few individuals.

Aktipis (2006) used a rather crude memory model. Might a more nuanced one, in which memory performance is responsive to statistical regularities of past social encounters, suggest ways in which heuristics can exploit memory? As an example, consider the culture and ritual associated with the buying of rounds in an English pub (Fox, 1996). Essentially, in the latter case, drinks are purchased in rounds, where one member of the group buys the round, but the expectation is that the other members will pick up a round later that evening or farther in the future. Suppose you enter a bar filled with beer drinkers and must decide how to optimize your round buying. Your only concern is your future consumption of drinks, so all you want is to find a person whom you will likely meet again. Basically, you need to think about how long is the period in the future L in which you have a reasonable chance of seeing your bar mate again. This could be over your lifetime (e.g., if you are a regular) or just a few days (e.g., if you are at a conference). What is the probability you will see her again in the future? Suppose it is Monday and you will be leaving town on Wednesday morning. Then, in this simplest case, where $L = 1$, $p_{L=1} = p_{Tue}$, that is, the relevant probability consists only of the probability that you will see her on Tuesday. If $L = 2$ (i.e., you will be leaving on Thursday morning), then the relevant probability is the sum of the probability that you will see her on Tuesday, and the probability that you will see her on Wednesday, but not on Tuesday: $p_{L=2} = p_{Tue} + p_{Wed}(1 - p_{Tue})$. Generalizing this to periods of length L, yields:

$$P_L = \sum_{i=1}^{L} p_i \prod_{j=1}^{i-1} \left(1 - p_j\right) . \qquad (1)$$

P_L is thus the probability of encountering the person at least once sometime during period L. As we have shown in our analyses of social contact, the probability p of having contact with someone will not be constant, but vary as a function of our history of contacts with that person. These probabilities will depend, on a first approximation, on how frequently and recently we have seen someone, as shown in Figures 2a and 2b. But is it reasonable to assume that a person, or any animal for that matter, could calculate, or even approximate, P_L? Probably not. In the next section, we explore how to exploit the memory system, envisioned in the ACT-R cognitive architecture, to achieve something that approximates the calculation in Equation 1.

Using memory activation to predict future encounters. The ACT-R research program (Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004; Anderson & Lebiere, 1998) strives to develop an encompassing theory of cognition, specified to such a degree that phenomena from perceptual search to the learning of algebra might be modeled within the same framework. The core of ACT-R is constituted by a procedural system, consisting of rules that describe the cognitive steps needed to carry out a given procedure, and a declarative memory system for facts. The declarative memory system consists of records that represent information (e.g., about the outside world, about oneself, etc.). These records take on activations that predict whether information will be needed to achieve a processing goal, such as recalling the name of someone you have just encountered at a conference. Activation is reflected in the accessibility of memory records, that is, whether and how quickly they can be retrieved. In other words, these activations reflect the "need probability"—the probability of needing a particular piece of information—from a rational analysis of memory (e.g., Anderson & Milson, 1989), and thus the kinds of social contact statistics we found are quite relevant.

Activation, A, corresponds to the log odds (i.e., $\ln[p/(1-p)]$) that the record in memory will be needed to achieve a processing goal. A record's activation is calculated as the sum of the base-level strength of the record and the activation that the record receives from the current context. For our purpose, however, we ignore context, and activation is thus simply the base rate activation rooted in the environmental pattern of occurrence.[8] Specifically, A_i is determined by an equation that strengthens and decays a record's activation according to the pattern with which it has been used.

$$A_i = \ln\left(\sum_{k=1}^{n} t_k^{-d}\right) = \ln\left(\frac{p}{1-p}\right), \quad (2)$$

where the record has been encountered n times in the past at lags of $t_1, t_2, \ldots, t_n$. . Finally, d is a decay parameter that captures the amount of forgetting in declarative memory, thus determining how much information about an item's environmental frequency is retained in memory. Figure 4a illustrates how activation builds and decays. Suppose that you see bar mate Norm[9] on January 6th and not again until January 11th. In this case the activation of the memory record for Norm would decay away in the interim. From the 11th on, the activation contributed by the second meeting would be added to the activation remaining from the first meeting, to give the total activation.[10]

How would memory of social contacts be represented in this system? And how could the calculation of P_L (i.e., the probability that one will have contact with a particular person in the future) be derived directly from the activation of a memory trace about a social contact? Keeping in mind that A is the log-odds of needing a record of having had previous contact with someone, and that $p = o/(1+o)$, then the probability of needing a record, p, is

$$p = \frac{e^A}{1+e^A} \quad (3)$$

and so,

$$1-p = \frac{1}{1+e^A} \quad (4)$$

Substituting Equations 3 and 4 into Equation 1 yields,

$$P_L = \sum_{i=1}^{L} \frac{e^{A_i}}{1+e^{A_i}} \prod_{j=1}^{i-1} \frac{1}{1+e^{A_j}} \ . \quad (5)$$

Equation 5 suggests that if it were possible to project activation into the future, we could use activation of a memory record of someone to estimate the probability of contact with that person during the next L time units, P_L. Recall that this probability is critical for whether an individual should play TFT in a prisoner's dilemma. Given that the system does not have direct access to a calculation, such as Equation 6, one possibility would be to use

[8] Although we ignore contextual information in this chapter, here is how it could enter the picture. If, for example, Friday is Cliff's regular day to go to the pub, information about the day of the week could be used to boost slightly information associated with Norm, thus making activation sensible to context information. Note that this would reflect the weekly cycles apparent in Figure 2b.

[9] Norm and Cliff were bar flies in the popular television series *Cheers*.

[10] In ACT-R, it is not possible to access sub-symbolic quantities like activation directly. Rather, these must be gauged by whether and how quickly a memory record can be retrieved (Hertwig, Herzog, Schooler, & Reimer, 2008; Schooler & Hertwig, 2005). For our purposes, we will suppose that there is some transformation of activation, A, into a sensation of fluency or familiarity, such that if $A_i > A_j$ then, $f(A_i) > f(A_j)$.

current activation as proxy for future activation. Although such a strategy might often be successful, Figure 4b illustrates how it can lead to systematic errors. Suppose that you meet bar mate Cliff for the first time on January 26th, and then you meet both Norm and Cliff on the 28th. On the 28th, Cliff will be more familiar ($A_{Cliff} = 0.53$) than Norm ($A_{Norm} = 0.44$), so you would buy Cliff a beer, assuming d is set to ACT-R's default value of .5. However, projecting out into the future that activation for each suggests that you are more likely to see Norm than Cliff again. Cross-overs like this are also reflected in the statistics of Figure 2c (which were due to differences in the spacing of contacts) and represent a strong regularity in the memory literature known as Jost's Law (Jost, 1897): If two memories are of equal accessibility now, the older memory will be relatively more accessible in the future.

This analysis shows that if we assume that the familiarity experienced with some person represents some function of memory activation, this familiarity may not be a reliable indicator of the probability of future encounters. You may be more likely to encounter someone again in the future who is less familiar now than someone else who is more familiar. Even though the memory system (at least within in the ACT-R framework) has the needed information encoded, familiarity alone does not suffice to reliably predict the probability of future contact. Adding only a little information beyond familiarity, however, may help. For instance, if two people are equally familiar to a person but she remembers that the last contact for one of them was more distant in the past than for the other, she could correctly predict that her probability of contact is higher for the person with the less recent last contact. Nevertheless, we do not mean to imply that using familiarity as the basis of heuristics to predict future encounters could not be effective. How well such a heuristic will work depends on details both about the statistics of the social environment and how memory works. For example, familiarity might be a reliable indicator of probability of contactin the short run (i.e., small L) but not for probability of contact in the long run (i.e., large L). It is currently unclear what mechanisms people actually use to judge the probability of future contact with a particular person. Our analyses, however, might help to identify these mechanisms. If people relied primarily on familiarity to make that judge contact probabilities, then their judgments would display predictable, systematic errors. If people steered clear of familiarity-based errors, this would suggest that they incorporate additional cues, such as recency or frequency information, in their judgments.

In sum, the probability of encountering a particular network member again in the future represents key information for deciding whether to cooperate with her or not. This probability is more or less encoded in the relative activation of memory records formed by past encounters with the person. Memory thus offers a valuable resource to be exploited for making key social decisions. The catch is that predicting, based on activation alone, which person you are more likely to encounter in the future may, in predictable ways, lead you astray.

Conclusions and Outlook

"We learn that the target person is not surrounded by acquaintance points, each of which is equally likely to feed into an outside contact; rather, there appear to be highly popular channels for the transmission of the chain". Thus concludes Stanley Milgram (1967, p. 66) from an often overlooked result of his classic study of the small-world phenomenon. In this chapter, we reported evidence that a skewed distribution in the usage of a person's contact channels characterizes social contact more generally. We live in a small world not only because we can reach a large number of people in a small number of steps, but also our daily encounters are concentrated on a very small number of people. We have argued that this skewed pattern in the probability of contact is particularly relevant for the evolution of cooperation. Most existing models of cooperation ignore contact patterns, but our results suggest that they may have strong influences on cooperation. In particular, the evolution of the simple heuristic TFT may be quite limited when the probability of interaction is skewed.

Further, we found systematic relationships between aspects of past encounters and the probability of future encounters. These patterns mirror results found in other forms of social interaction (word usage, email contact), and, intriguingly, they show characteristics that have also been found in patterns of memory performance. Based on this correspondence, we have highlighted ways in which social heuristics might exploit memory to organize the distribution of cooperative behavior, as well as the constraints associated with such social heuristics.

Social contacts provide the infrastructure for social exchange and social transmission. Our results should therefore be relevant for a number of areas beyond the evolution of cooperation. For instance, the influence of connection strength and network contact dynamics has implications for models of the transmission of viruses, fads, and social information. Like cooperation, the study of cultural transmission focuses mainly on the network structure, neglecting the distribution of contacts. Yet the strength of connections may have important implications for transmission (Read, Eames, & Edmunds, 2008). Concerning the role of memory for predicting the probability of future contact, our results highlight the importance of testing the accuracy of memory to predict future social contact, as well as the relative importance of the various aspects of past contacts—frequency, recency, and spacing—in predicting the chance of meeting someone again. In short, from patterns of memory to social interactions, the past is a powerful predictor of the future and determinant of the present.

References

Amaral, L. A. N., Scala, A., Barthélémy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences (USA), 97*, 11149-11152.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglas, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1030-1060.

Anderson, J. R., & Lebiere, C. (1998). *Atomic components of thought.* Hillsdale, N.J.: Erlbaum.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*, 703-719.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*, 396-408.

Aktipis, C. A. (2006). Recognition memory and the evolution of cooperation: How simple strategies succeed in an agent-based world. *Adaptive Behavior, 14*, 239-247.

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books, Inc.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*, 1390-1396.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks, *Science, 286*, 509-512.

Brandon, R. N. (1996). *Concepts and methods in evolutionary biology*. Cambridge University Press.

Cepeda, N., Pashler, H., Vul, E., Wixted, J., & Rohrer, D. (2006). Distribution of practice in verbal recall tasks: a review and quantitative synthesis. *Psychological Bulletin*, *132*, 354-380.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (in press). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science*.

Copher, J., Kanfer, A., & Walker, M. B. (2002). Everyday communication patterns of heavy and light email users. In B. Wellman & C. Haythornthwaite (Eds.), *The Internet in everyday life* (pp. 263-288). Oxford, UK: Blackwell.

Dodds, P. S., Muhamad, R., & Watts, D. J. (2003). An experimental study of search in global social networks. *Science*, *301*, 827-829.

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *22*, 469-493.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, *6*, 178-190.

Fox, K. (1996). *Passport to the pub: The tourist's guide to pub etiquette.* London: Brewers and Licensed Retailers Association. http://www.sirc.org/publik/passport.pdf

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, *15*, 1-16.

Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1191-1206.

Hill, R., & Dunbar, R. (2003). Social network size in humans. *Human Nature*, 14, 1, 53-72.

Jones, B., & Rachlin, H. (2006). Social discounting. *Psychological Science, 17*, 283-286.

Jost, A. (1897). Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen [The strength of associations in their dependence on the distribution of repetitions]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *16*, 436-472.

Killworth, P. D., Bernard, R., & McCarthy, C. (1984) Measuring patterns of acquaintanceship. *Current Anthropology*, *25*, 381-397.

Kleinfeld, J. (2002). The small world problem, *Society, 39,* 61-66.

Kraines, D., & Kraines, V. (1989). Pavlov and the prisoner's dilemma. *Theory and Decision, 26*, 47-79.

Laland, K. N., Odling-Smee, F. J., & Feldman, M.W. (2000). Niche construction, biological evolution and cultural change. *Behavioral and Brain Sciences, 23,* 131-175.

Lewontin, R. C. (1983). Gene, organism and environment. In D. S. Bendall, *Evolution from molecules to men* (pp. 273-285). Cambridge, UK: Cambridge University Press.

Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E., & Aberg, Y. (2001). The web of human sexual contacts. *Nature, 411*, 907-908.

Masuda, N., & Aihara, K. (2003). Spatial prisoner's dilemma optimally played in small-world networks. *Physics Letters A, 313*, 55-61.

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.

Milgram, S. (1967). The small world problem. *Psychology Today*, *2*, 60-67.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review, 45*, 167-256.

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics, 46 (5),* 323-351.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*, 1560-1563.

Nowak, M. A., & May, R. M. (1992). Evolutionary games and spatial chaos. *Nature, 359*, 826-829.

Nowak, M. A., & Sigmund, K. (1993). A strategy of win–stay, lose–shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature, 364*, 56-58.

Ohtsuki, H., Hauert, C., Lieberman, E., & Nowak, M. A. (2006). A simple rule for the evolution of cooperation on graphs and social networks. *Nature, 441*, 502-505.

Pachur, T., Schooler, L. J., & Stevens, J. R. (2010). We'll meet again: Revealing distributional and temporal patterns of social contact. Manuscript submitted for publication.

Pollock, G. B. (1989). Evolutionary stability of reciprocity in a viscous lattice. *Social Networks, 11*, 175-212.
Price, D. J. D. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, *27*, 292-306.
Rapoport, A., & Chammah, A. N. (1965). *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor, MI: University of Michigan Press.
Read, J. M., Eames, K. T. D., & Edmunds, W. J. (2008). Dynamic social networks and the implications for the spread of infectious disease. *Royal Society Interface*, *5*, 1001-1007.
Santos, F. C., Rodrigues, J. F., & Pacheco, J. M. (2006). Graph topology plays a determinant role in the evolution of cooperation. *Proceedings of the Royal Society of London, Series B, 273*, 51-55.
Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, *112*, 610-628.
Schooler, L. J., Serio-Silva, J. C., & Rhine, R. (2010). *Reflections of early hominid environments in human memory: Studies of primate ranging behavior.* Manuscript in preparation.
Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika*, *42*, 425-440.
Stevens, J. R., & Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences, 8*, 60-65.
Travers, J., & Milgram, S. (1969). An experimental study of the small world problem. *Sociometry*, *32*, 425-443.
Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35-57.
Watts, D. J. (1999). *Small worlds: The dynamics of networks between order and randomness*. Princeton, NJ: Princeton University Press.
Watts, D. J. (2003). *Six degrees: The science of a connected age.* New York: Norton.
Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*, 440-442
Zhou, W.-X., Sornette, D., Hill, R. A., & Dunbar, R. I. M. (2005). Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B*, *272*, 439-444.

Tables

Table 1. Interaction patterns and cooperation in the iterated prisoner's dilemma.

Pattern of interaction	Definition	Stability of cooperation
Network structure		
Regular network	Agents exist on a spatial network in which they interact with a fixed neighborhood of partners.	Tit-for-tat (TFT) is stable with fixed partners[1], and unconditional cooperation can evolve[2].
Small-world network	Agents exist on a regular network in which a few connections are rewired outside of the local neighborhood.	A few random connections can support cooperation, but many random connections inhibit cooperation[3,4].
Scale-free network	Agents differ in the number of partners with whom they interact.	Unconditional cooperation can evolve, due to both growth and preferential attachment[5].
Probability of contact		
Equal	Agents interact equally often with all partners, either by playing repeatedly against the same partner or by playing with a randomly chosen partner.	Unconditional cooperation is not stable, but TFT may succeed with very low error rate[6].
Skewed	Agents interact with different partners at different frequencies.	We do not know how a heterogeneous probability of contact will influence cooperation, but it likely will make unconditional cooperation and TFT difficult to implement.

Notes: [1]Pollock, 1989; [2]Nowak & May, 1992; [3]Watts, 1999; [4]Masuda & Aihara, 2003; [5]Santos, Rodrigues, & Pacheco, 2005; [6]Axelrod & Hamilton, 1981.

Figures

Figure 14-1. Distribution of contacts across network members in the analysis of TP's social contacts. (a) The probability of contact with a particular network member on a given day is highly skewed, with the large majority of members having a probability of less than 0.1. (b) When the network members are ordered according to contact probability, the relationship between rank and contact probability resembles a power-law distribution. (c) The pattern of social contacts with various network members across a period of 739 days varies, some showing a uniform or spaced pattern and other showing a clumped or massed pattern.

Figure 14-2. Regularities in social contact. (a) Probability of contact *p* is a linear function of the frequency of past contacts *f*. (b-c) The odds of contact *o* is a power function of the number of days since the last contact, *r*, appearing as a straight line on a log-log plot; even (as shown in c) when *f* is taken into account. (d) *p* (shown as running average using a five-day bin size) depends on both the number of days since the last contact and whether the past contact was massed or spaced, holding *f* constant at two. The R^2s in (b) and (c) refer to the fit of the log-transformed power (and thus linear) functions to the log-transformed data.

		Partner's strategy	
		Cooperate	*Defect*
Player's strategy	*Cooperate*	Reward	Sucker's pay-off
	Defect	Temptation to cheat	Punishment

Figure 14-3: Payoff matrix for the prisoner's dilemma. The player can choose to cooperate or defect (row strategies) at the same time that the partner chooses to cooperate or defect (column strategy). If both cooperate, the player receives R, the reward for cooperating. If both defect, the player receives P, the punishment for defection. If the player cooperates and the partner defects, the player receives S, the sucker's payoff, and if the reverse is true, the player receives T, the temptation to cheat. A prisoner's dilemma occurs if $T > R > P > S$. The best strategy is always to defect because, regardless of the partner's strategy, defection yields a higher payoff than cooperation, regardless of whether the partner cooperates ($T > R$) or defects ($P > S$).

(a)

(b)

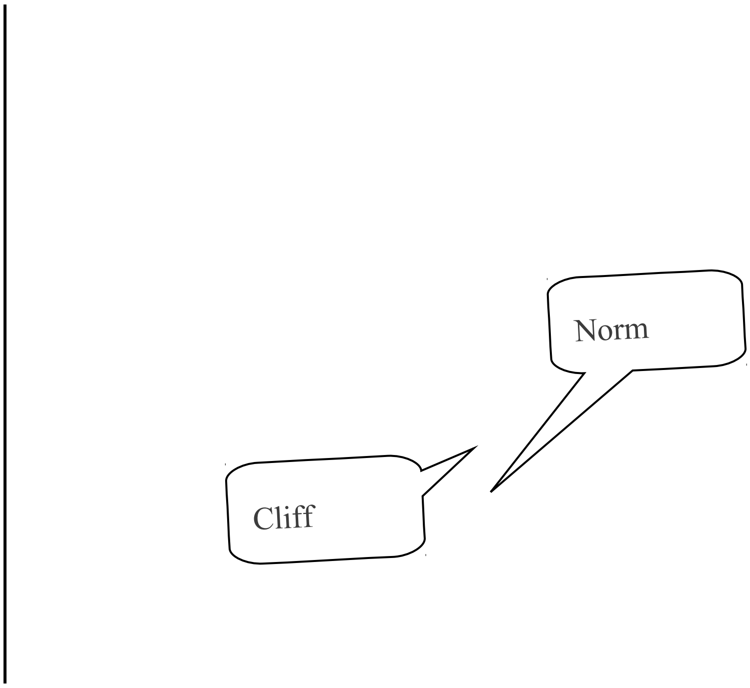

Figure 14-4: (a) Each encounter increases the activation of a memory record. The activation of the individual encounters is represented by the thin black lines. The activation of the sum of the two encounters is shown by the line punctuated by black dots. (b) Current levels of activation do not perfectly predict future levels. On the 28th, Cliff will be more familiar than Norm, but Norm's higher future activation in March suggests that you will be more likely to meet Norm in March than Cliff.