Neuronal network analyses: premises, promises and uncertainties

David Parker*

Department of Physiology, Development, and Neuroscience, University of Cambridge, Cambridge, UK

Neuronal networks assemble the cellular components needed for sensory, motor and cognitive functions. Any rational intervention in the nervous system will thus require an understanding of network function. Obtaining this understanding is widely considered to be one of the major tasks facing neuroscience today. Network analyses have been performed for some years in relatively simple systems. In addition to the direct insights these systems have provided, they also illustrate some of the difficulties of understanding network function. Nevertheless, in more complex systems (including human), claims are made that the cellular bases of behaviour are, or will shortly be, understood. While the discussion is necessarily limited, this issue will examine these claims and highlight some traditional and novel aspects of network analyses and their difficulties. This introduction discusses the criteria that need to be satisfied for network understanding, and how they relate to traditional and novel approaches being applied to addressing network function.

Keywords: neuronal network; network; microcircuit; brain; neural circuit; behaviour

1. ‘LET US BE REALISTIC: WE MUST INSIST ON THE IMPOSSIBLE’. CHE GUEVARA

There are obvious reasons for wanting to understand the brain: one is technological, to apply principles of nervous system function to computing and robotics: a second is of course is clinical, to provide rational treatments for neurological or psychiatric disorders. We know a lot about the cellular properties of the nervous system and continue to identify molecular, developmental and functional properties of neurons and synapses. At the opposite end of the scale, we can characterize and quantify behaviours and correlate them with activity imaged with increasing sophistication in different regions of the brain. However, between these two levels there is an ‘explanatory gap’ that has prevented us from explaining behaviours directly in terms of their underlying cellular and synaptic mechanisms (Dudai 2004; Parker 2006a). We are thus data-rich but lack knowledge of how to integrate these data into a coherent picture of brain function. This is evidenced by the limited extent to which we can intervene to correct aberrant functions: claims for effective interventions have been called the ‘lobotomy attitude’ (Dudai 2004) to illustrate that our current knowledge makes effective interventions unlikely in the foreseeable future. However, these claims exist, and they go beyond traditional roles in neurology and psychiatry to educational and legal practices and the enhancement of normal functions (Farah et al. 2004; Lynch 2004; Farah 2005; Goswami 2006; Parker 2006b). Barring ‘penicillin moments’ successful applications will require rational approaches based on genuine understanding of the cellular basis of nervous system function. It is not simply a matter of being right, but also of highlighting what we can genuinely claim to understand of the mechanisms underlying cognitive functions and behaviours. Failure to do this stifles endeavour, and could lead to interventions that at best will only fail.

Closing the explanatory gap between the cellular and behavioural levels ultimately requires knowledge of the neuronal networks that assemble the cellular components underlying sensory, motor and cognitive functions. These networks form the link in any understanding of how cellular properties influence behaviours, and their analysis is widely considered to be one of the major tasks facing neuroscience. Some claim that this understanding is imminent, to the extent that cognition and mind will soon be reduced to molecular and cellular mechanisms (the ‘reductionist epiphany’; Bickle 2007; see Markram 2006; Olsen & Wilson 2008; Sandberg & Bostrom 2008), while in other cases, the desire to have solved this question has led to claims that this understanding has already been achieved. These claims are difficult to evaluate without detailed knowledge of the specific systems studied. However, they often lack direct causal links and instead rely on assumptions and extrapolations between the molecular, cellular, network and behavioural levels (see Dudai 2004; Parker 2006a).

I have outlined this for the model network that I work on, the lamprey spinal cord locomotor network (see Parker 2006a). Repeated claims that this network
has been experimentally characterized (e.g. Grillner et al. 2005) to the extent that behaviour can be explained in terms of specific functional interactions between identified neurons (‘we can actually make the bridge from the molecular to the behavioural level’; Grillner 1999) are very widely accepted within and outside the field. However, these claims and their acceptance ignore the uncertainties over the characterization of network components (see Rovainen 1983; Buchanan 1999; Parker 2006a) and hide very significant gaps in our understanding of the basic network connectivity (figure 1; see Parker 2006a). These uncertainties alone (there are more) logically preclude any claims of direct deterministic links between cellular properties and network function and behaviour (see Parker 2006a; figure 1). A similar critique can be made of the network controlling the gill withdrawal reflex in *Aplysia*. The claimed deterministic links between molecular and cellular mechanisms underlying this behaviour arguably provide the archetypal model of a characterized system (Kandel 2001). However, these claims are again associated with uncertainties over the proposed cellular mechanisms (see Glanzman 2007) and the network components examined (Zecevic et al. 1989; Trudeau & Castellucci 1993, 1995; Hickie et al. 1997), which again weaken any claims to deterministic links. The characterizations in these and similar cases ultimately suffer from various logical fallacies (it is ironic that ‘logic’ has become a keyword in the network field), and fail to provide evidence sufficient in number, kind or weight to support the claims that are being made. The focus on simplification and synthesis means that attention is generally wasted effort and reduced motivation to address the uncertainties and false claims (which may not be correct for a characterized network. The excitatory input (E on this diagram) to these cells is also problematic. Repeated claims, most recently by Grillner & Jessell (2009), state that they ‘excite all types of spinal neurons’, but there is no demonstrated connection between the E and the ScINs, making the claim misleading. Grillner & Jessell go on to say that ‘in lamprey and tadpole intrinsic synaptic excitation within pools of the excitatory premotor interneurons … account for the burst generation in combination with their membrane properties’, and ‘gap junctional connectivity has been reported in this interneuron pool’. Neither is established for lamprey: although it seems likely that the evidence for direct connectivity between the EINs (Parker & Grillner 2000) will support bursting, these interactions are very poorly understood, and there is no evidence for gap junctional connectivity. Detailed information is, however, available on excitatory network interneurons is available, however, from a series of detailed analyses in the tadpole (Li et al. 2006; Roberts et al. 2008, in press). For crossing inputs in the lamprey network the situation is worse: claims that they ‘cross the midline to inhibit all neuron types on the contralateral side’ (Grillner & Jessell 2009) are not justified. No CC interneuron inputs to E cells are known, and ScIN connectivity is essentially unknown (they are only known to inhibit motor neurons), despite claims that they have been determined experimentally (Grillner et al. 2005). In reality there are significant gaps in our knowledge simply at the level of the organization of this network, with obvious implications for any attempt at a functional explanation of how the network output is generated. The diagram also omits ipsilateral inhibitory inputs and crossing excitatory inputs (see Parker 2006a for a more detailed discussion). The apparent characterization here is achieved by assumption and extrapolation and by omitting reference to highlighted gaps in our knowledge (Rovainen 1983; Buchanan 1999; Parker 2006a).

This issue will critically discuss neuronal network analyses. It will look at the extent to which we understand 'simpler' model networks: it is often forgotten that these systems proved to be more complex than initially thought. In addition to the significant insights into the principles of network operation that they have provided, these systems serve as a historical record of potential sources of error and limits to understanding that should inform current approaches and claims (these aspects are highlighted by Allen Selverston (Selverston 2010)). The issue will also examine the utility of certain experimental approaches (Astrid Prinz (Prinz 2010) and Smeal et al. (2010) on theoretical and computational approaches; Patrick Whelan (Whelan 2010) on molecular approaches), and features that may extend the traditional criteria needed to understand networks (Vladimir Brezina

---

*Phil. Trans. R. Soc. B* (2010)
2. CRITERIA FOR NETWORK UNDERSTANDING

Several factors have to be considered in any network analysis. Some of these are trivialities, but can nevertheless add significant confusion. For example, in addition to neuronal networks (i.e. interactions between functionally related groups of neurons), the term network can also be used in different contexts, from genetic networks and molecular interactions to interactions between brain regions. Confusion can also result from the variable terms that can be used to refer to neuronal networks (e.g. CPG, circuit, network, assembly, microcircuit and oscillator). The introduction of new terms (microcircuit is the most recent) can give the impression of some novelty or advance when it is just a renaming: the new term is seldom explained and the factors that distinguish it from previous terms are often unclear. Finally, confusion can also result from the tendency to make general definitions based on the specifics of particular networks (e.g. ‘network interactions are fundamentally inhibitory’; Yuste et al. 2005).

A more difficult question is to define what we mean by understanding, and how we would know that a network was understood (i.e. that there was nothing else to address). Understanding can reflect several aspects, but would imply the ability to explain a networks output in terms of the spatial and temporal organization of its cellular and synaptic components; to predict the response of the network to inputs or perturbations; and to understand the functional role of network activity. In this context modelling is essential, either acting as a heuristic to generate hypotheses that direct experimental analyses, or to test if available experimental data can reproduce or explain network function. Limits to experimental data mean that assumptions are inevitable. But even where data are available, the complexity of detailed models can make them as difficult to understand as the actual system (e.g. Greenberg & Manor 2005), and even relatively simple questions in artificial neural nets are practically non-solvable (NP-complete; Gu et al. 2009). Models are thus inevitably simplified and cannot be refuted in assumptions that they reflect actual systems. Nevertheless, the insight gained can give further direction rather than being lost in a mess of detail, particularly when the assumptions of the model are explicitly stated.

The criteria for understanding a neuronal network have been outlined several times (e.g. Selverston 1980; Getting 1989; Yuste 2008). The minimal criteria are the characterization of the network organization (its component neurons and direct synaptic connections); the identification of the functional properties of the cellular and synaptic components; and determining how these properties directly influence the flow of activity through a network, and thus the network output and associated behaviours. Identifying network neurons is the fundamental first step, but even this basic requirement can be difficult. Unless neurons belonging to a particular network are strictly segregated, it cannot be assumed that all neurons in a particular region belong to the same network or serve a particular function. This makes the problem far more than simply classifying the neuronal subtypes present in a region of the nervous system. Traditional criteria for identifying a network neuron are that the neuron is active when the network is active and that it influences the network output. However, both criteria can lead to the erroneous exclusion or inclusion of neurons. For example, an active neuron might not be a component of the active network but may instead be activated downstream (this reflects the functional limits drawn around the network, a potentially contentious question in itself). Conversely, the absence of correlated activity does not rule a neuron out as a network component as a cell may only be active during certain types of network activity: to avoid errors of exclusion cells must be examined over a range of physiologically or behaviourally relevant activity patterns (assuming that these are known and can be evoked in experimental preparations). Caveats also apply to the second criterion that the activity in a type of neuron influences the network output. Activating or inhibiting a small proportion of neurons in a large population may not affect the network output (but see Bonifazi et al. 2009) leading to erroneous exclusion; conversely, activity in non-network neurons could alter the output (erroneous inclusion). Even if all neurons in a population were activated or silenced, degeneracy of redundancy of function (Tononi et al. 1999) or compensatory adaptations (which can be rapid; Frank et al. 2006) could further complicate interpretations.

Once network neurons have been identified, the network organization needs to be characterized. This demands structural and functional evidence for monosynaptic (direct) connections between identified network neurons. Caution also has to be exercised here: the history of network analyses again highlights many errors and potential pitfalls (see Berry & Pentreath 1976). These occurred using approaches that are currently being used routinely. The typical physiological criterion for a monosynaptic connection is the latency of the postsynaptic response after presynaptic stimulation. For example, Olsen & Wilson (2008) state ‘if one can demonstrate that a precisely timed depolarization of one neuron evokes a short-latency synaptic response in the other neuron, then a direct connection is unequivocal’ (my italic). While it may be indicative, latency is far from unequivocal, and could actually provide one of the weakest criteria for a monosynaptic connection (see Berry & Pentreath 1976). Its use rests on assumptions of synaptic delays and axonal conduction velocity, properties that can vary widely. Better physiological criteria for a
monosynaptic input are that it occurs reliably and with constant latency during high frequency stimulation trains; it persists in high divalent cation solution (which raises the spike threshold through surface screening and can abolish polysynaptic inputs by silencing neurons in polysynaptic pathways); and that it is gradually reduced in low calcium Ringer (a polysynaptic connection will again drop out causing a sudden failure rather than a gradual reduction of the postsynaptic input). However, no approach gives unequivocal answers and interpretations can be complicated by synaptic arrangements (e.g. dendro-dendritic interactions, graded or electrical synapses). These criteria and approaches also only apply to analyses using paired electrophysiological recordings; they lose predictive power with extracellular stimulation or recording, and optical imaging techniques do not currently have sufficient temporal resolution to determine monosynaptic over polysynaptic connections (see below). Defining monosynaptic inputs ultimately requires a combination of a range of physiological and ultrastructural analyses (see below).

From the identification of network neurons and the network architecture the functional properties of component cells and synapses must be determined. Space prevents even a brief list of these properties. For example, it is estimated that more than 200 substances can act as transmitter substances (Thomas 2006), while the voltage-dependent ion channels that determine resting and active cellular properties form a superfamily of at least 143 genes (Yu et al. 2006), with further diversity in channel and cellular function resulting from alternative splicing, post-translational modification, and varying combinations of channel subunits (Gutman et al. 2005). The particular complement of cellular and synaptic properties present determines the functional signature of cells and their contribution to the patterning of network activity. These properties thus cannot be generalized but must be characterized in specific classes of network neurons and synapses. This introduces a very significant burden on network analyses, as the wide range of cellular and synaptic properties means that the functional state of a neuron can take many values: even assuming just two values, active or inactive, the number of states in a system containing \( n \) components would be \( 2^n \), making the complete functional description of even moderately sized networks far more difficult than a description of the network architecture alone. While few people would probably doubt the importance of functional properties, their analysis nevertheless seems to be underappreciated. For example, after a very extensive discussion of the importance of functional properties, their analysis and ultrastructural analyses (see below). Defining monosynaptic inputs ultimately requires a combination of a range of physiological and ultrastructural analyses (see below).

... defined in many recent review issues (e.g. Selverston 1980; Getting 1989), but still seems to need repeating. Overly positive claims tend to attract the attention of those on the periphery or new to a field, while the analytical and conceptual gaps and problems receive less attention. In terms of network analyses, electrophysiology currently remains the dominant approach. However, it has significant limitations: while extracellular recording can readily assay activity from single cells or cell populations with high spatial and temporal resolution, it cannot identify cellular and synaptic properties or subthreshold events; conversely intracellular or whole cell recording allows the properties of individual or a limited number of single cells or synapses to be examined in great detail, but scaling these properties up to the population or network level (e.g. the spatio-temporal patterning and propagation of activity) is difficult (e.g. Gervais et al. 2007). Novel techniques based on molecular biology, physics and engineering that aim to overcome the limitations of traditional electrophysiological approaches are actively being sought (see Miyakawa & Schnitzer 2007; Scanziani & Hausser 2009). These techniques have been highlighted in several recent review issues (Current Opinion in Neurobiology 2007 17(5); Science (2009) 326; Nature (2009) 461, 7266). They will be discussed briefly here in the context of the requirements for network understanding.

First, the application of molecular genetic techniques has provided several important approaches to...
network analyses. Genetically expressed fluorescent or viral markers in specific cell types facilitate analyses of network organization, and have provided significant insights into the morphology, organization and development of network neurons and connections. Genetically encoded reporters can also be used to monitor changes in gene expression as a result of neural activity (see Barth 2007). The use of molecular techniques in functional network analyses traditionally depends on the loss of function (i.e. knocking out or silencing identified or suspected network neurons), and has been claimed to allow the ‘dissection’ of neuronal networks (Kiehn & Kullander 2004; Luo et al. 2008). While progress has been made in developing genetically encoded systems for the inducible and in some cases reversible silencing of network components, none is ideal and the links between knock-outs and network function can be simplistic. Further optimization of these techniques is needed to reduce heterogeneous effects and allow reproducible responses to be targeted to specific changes in specific classes of network neurons (i.e. to overcome the effects of overlapping patterns of gene expression between cells and anatomical regions, and to prevent compensatory changes during development). Much is made of the ability to target specific cell types, but what is meant by a cell type? Specific classes of cells can be targeted based on the presence of particular markers, but whether these markers separate distinct functional classes of a cell is at best uncertain and often knowingly not the case (Gordon & Whelan 2006; O’Connor et al. 2009). Combinatorial strategies, where the transgene expression depends on the presence of two or more promoters, offer the hope of greater spatial precision (Luan & White 2007). Greater temporal precision is also needed to allow more rapid inducible effects in mature systems to circumvent compensatory responses to the manipulations: even the fastest inducible effects (e.g. allatostatin) currently exceed the time needed for potential compensatory adjustments or plasticity (Frank et al. 2006), with obvious complications for functional interpretations. Ideally a list of cell-specific gene expression patterns could be consulted for all cell types in a network and used to knock out or perturb the function of only the cells of interest. But even if a pure population of cells were affected (and assuming no functionally relevant heterogeneity in the functional properties of these cells or in expression levels of probes or targeted factors), many of the caveats outlined above for defining network neurons and monosynaptic connections would still apply. Thus, the presence of an effect could only identify (with several caveats) some necessity for the component but not its actual functional role, and all outcomes (including the absence of an effect) are subject to complications introduced by the degeneracy, redundancy or compensation of function (Tononi et al. 1999; Marder & Goaillard 2006). Even in the ideal situation (i.e. specific manipulation of a single functionally defined class of cell), heterogeneity within single populations could introduce a significant complication. This can only be dismissed if the heterogeneity was assumed to be of no functional importance, a view that is gradually being challenged (see below; Soltesz 2006). Patrick Whelan (Whelan 2010) will discuss these issues further. A second recent technique is high-throughput ultrastructural analyses (e.g. serial block face scanning electron microscopy (Denk & Horstmann 2004); array tomography coupled with fluorescent imaging (Micheva & Smith 2007)). These techniques aim to map network neurons and synaptic connections to provide a complete anatomical description of networks (‘connectomics’; Lichtman & Sanes 2008) either at the microscopic or mesoscopic scales (Bohland et al. 2009). Proponents of connectomics claim that it is controversial because the difficulty of parsing connections between neurons at the turn of the twentieth century lead to a focus on individual cells and synapses at the expense of network analyses (see Lichtman & Sanes 2008). This is clearly not the case: the basic aims of connectomics (i.e. characterising network architectures) have been a cornerstone of network analyses from the start. However, controversy does arise from connectomist claims: these argue (correctly) for the importance of structural data, but incorrectly that structure can predict circuit function and synaptic efficacy (Lichtman & Sanes 2008; Sandberg & Bostrom 2008). Connectomists compare their critics to those of the Human Genome Project, and say by analogy that the criticisms will disappear as connectomics proceeds (Lichtman & Sanes 2008). This is likely to be true: criticisms of the Genome Project disappeared as the recognition of its exaggerated claims (e.g. that the sequence would explain function) led to the necessary move to transcriptomics, proteomics and metabolomics. This will be mirrored in connectomics: while structural data allow some inferences on network output, synaptic type, or transmitter content, these are very limited at best (see Peters & Palay 1996; Hökfelt et al. 2000). Examples abound: functional and non-functional (silent) synapses do not differ structurally (Atwood & Wojtowicz 1999); neuronal morphology may (Kasper et al. 1994) or may not (Chang & Luebke 2007) correlate with functional properties; a particular output can be generated by diverse network configurations (e.g. degeneracy; Tononi et al. 1999) and the same configuration can generate diverse outputs (see Elson et al. 2002; Prinz et al. 2004). Nevertheless, despite these caveats, connectomics will allow functional analyses to be performed in the context of detailed ultrastructural or anatomical arrangements. This is an important aspect that is often lacking in network analyses, and can influence (if not determine) the type of network output that is generated (Galan 2008; Bohland et al. 2009). A knowledge of the network organization at the mesoscopic or ultrastructural level could provide insights that facilitate the understanding of functional properties and processing (e.g. the structural basis for computational interactions in the cerebellar glomerulus (Hamann et al. 2002; Mitchell & Silver 2003), and electrical synapse-mediated inhibition (Korn & Faber 1976)). A complete anatomical reconstruction should thus be welcomed for any network (e.g. as exists for Caenorhabditis elegans, White et al. 1986; Chen et al. 2006), but it is not a trivial...
issue for most networks. The major bottleneck is the need for efficient segmentation strategies to determine the direct relationships between elements needed for the reconstruction of the examined region. This may be overcome by automated discrimination of axons, dendrites and synapses combined with multiple fluorescent labelling (Livet et al. 2007), possibly using machine-learning approaches, but this still needs to be demonstrated. Estimates for the complete ultrastructural reconstruction of networks vary from 2–3 years for a single cortical column (assuming improved automated segmentation strategies; Smith 2007) to decades for the whole visual cortex (Sandberg & Bostrom 2008), while the complete mesoscopic description of the mouse brain is estimated to take five years; Bohland et al. 2009).

An associated approach for investigating network connectivity is the engineered viral labelling of neurons using modified vectors that express fluorescent proteins (see Callway 2008). The problem here is to limit infection to allow the identification of monosynaptic over polysynaptic pathways (it again provides no direct insight into functional properties). The rate of viral infection is dependent on the strength of the connections (strong polysynaptic pathways will be labelled preferentially over weak monosynaptic connections), which could recapitulate some of the problems and errors of early electrophysiological analyses of synaptic connectivity (see Berry & Pentreath 1976). A recent novel strategy using modified viruses that lack the glycoprotein coat needed for retrograde viral transfection of presynaptic neurons was claimed to allow putative monosynaptic connections to be identified by preventing the virus spreading to cells further upstream (Wickersham et al. 2007). Mono-synaptic connections were verified in 9 of 11 cases where direct confirmation was sought electrophysiologically. While plausible reasons could exist for the cases where monosynaptic connections were not supported (e.g. that the neurons they were immediately presynaptic to had died), it begs the question whether it is reasonable to assume that failures to show monosynaptic inputs reflect technical or uncontrolled aberrations. Nevertheless, given the difficulties of determining of network organization, viral labelling seems a potentially effective tool for mapping network connectivity.

Finally, inferences from observation have been highlighted in the context of imaging using optical probes. Imaging again offers important tools (whether they are the ‘most important’ (Lichtman & Smith 2008) is a matter of debate). Observation is claimed to be the most efficient route to network understanding ‘due to the predominant importance of spatio-temporal dynamics’ (Lichtman & Smith 2008), a claim the authors support using the analogy that we learn the rules of football by watching. Similar analogies have been used previously (baseball (Kennedy 1971; Bullock 1980), an orchestra, (Buzsáki 2004)) to highlight how observation alone is a poor route to understanding. Again, an awareness of the history of network analyses would inform some of these claims. Imaging is not a new technique; it has been applied to network analyses for almost as long as electrophysiology (Cohen et al. 1978), but the techniques are becoming more sophisticated. While larger scale imaging (e.g. fMRI) cannot address neuronal network interactions (a voxel of several mm$^3$ could contain $5.5 \times 10^6$ neurons and $10^{10}$ synapses), single particle tracing aided by the use of bright quantum dots allows the movements of individual presynaptic and postsynaptic molecules to be followed (e.g. Zhang et al. 2009), and imaging of cell populations can better assess spatio-temporal aspects of network activity than single cell electrophysiology and can examine macroscopic network interactions with greater precision than is possible with extracellular recordings (Traub et al. 2008). Imaging also allows neuronal activity to be monitored non-invasively, avoiding potential complications of intracellular (input resistance drop caused by damage to the cell) and patch recordings (intracellular dialysis). However, imaging requires a reporter that converts neuronal responses into optical signals rather than providing direct measurements, it can perturb normal function (e.g. through chelation effects or membrane disturbances; Wallace et al. 2008; Akemann et al. 2009), and it currently lacks sufficient spatial resolution to image individual cells in networks, and sufficient temporal resolution to monitor single action potentials or synaptic inputs within neuronal populations (even temporal resolution of a few ms could confuse monosynaptic or polysynaptic pathways; Wang et al. 2007; Gradinaru et al. 2009). These limits are being addressed (e.g. Lillis et al. 2008), but the dream of using voltage sensitive dyes to monitor activity in large numbers of network neurons simultaneously at single cell resolution (that ‘the photon will replace the electron for probing neuronal function’; Scanziani & Hausser 2009) still seems far away.

A related advance is optical stimulation using genetically expressed channelrhodopsin and halorhodopsins (Nagel et al. 2003; Deisseroth et al. 2006; Zhang et al. 2007). This allows the rapid activation or inhibition of neurons with better temporal and spatial control than traditional electrical or chemical approaches (Petreanu et al. 2007; Sohal et al. 2009). Activating neurons using channelrhodopsin allows the impact of varying activity patterns in the cell to be examined, offering the significant advantage of gain of function analyses by activating anatomically or functionally defined cellular populations while examining the effect on the network output (Miesenböck 2009; Sohal et al. 2009). However, there is again an issue of limiting expression to defined populations, and of what is meant by ‘functionally defined’ cells. Does defined relate to cells with specific functional roles in a network, or is it simply to cells identified by some experimentally tractable property (e.g. the presence of a molecular marker for a particular cell)? unfortunately it is often the latter. Optical stimulation has been applied to attempts at mapping network connectivity (e.g. Arenkiel et al. 2007; Petreanu et al. 2007). While connectivity between regions can be examined, because it is currently not possible to selectively activate or label different populations a lack of connectivity only says that the labelled components are not connected and
thus is not definitive. This potential pitfall is illustrated by analyses of olfactory processing in *Drosophila*, where initial imaging studies (Ng et al. 2002) suggested a network organization that was not supported by subsequent electrophysiological analyses (Wilson et al. 2004). Analyses are also largely structural rather than functional: connectivity between regions can be identified, but analyses of the nature of the connections (e.g. the strength of inputs to different cortical layers; Petreanu et al. 2007) are complicated because individually labelled cells respond differently to stimulation as a result of uncontrolled differences in expression levels of the probes (Arenkiel et al. 2007; Petreanu et al. 2007; Sohal et al. 2009). In addition, while much is made of the ability of this technique to address the function in large populations, existing approaches can only drive action potentials in a small proportion of neurons (although this is still far greater than traditional electrophysiological approaches; Rickgauer & Tank 2009).

Technical advances that will address the limits of optical imaging and stimulation are being pursued. These will need to allow single cell resolution with sufficient spatial and temporal resolution for the optical recording of subthreshold and suprathreshold activity in identified neurons that are potentially deep in the tissue. This is necessary to avoid the current necessity of averaging multiple trials (Palmer & Stuart 2009), an approach that prevents spontaneous responses and variable properties from being examined. There is also a need for quantitative data on monosynaptic connections. The ability to perform these analyses will require faster and more efficient scanning (Lillis et al. 2008), greater depth penetration within densely packed neuropils (e.g. Ding et al. 2009), sub-millisecond temporal resolution to detect discrete cellular and synaptic signals, and the ability to repetitively activate neurons over physiologically relevant frequencies and time scales without desensitization of the probes. There is also a need to develop probes that overcome potential problems caused by the application of foreign agents into cells: calcium sensors can introduce chelating effects (Wallace et al. 2008), while conventional voltage sensitive dyes can generate toxic by-products and also add displacable charges to the membrane which can increase membrane capacitance and alter functional properties (Akemann et al. 2009). The main challenge to optical stimulation approaches is similar to that of molecular approaches to network analyses outlined above, namely the need for improvements in the specificity and spatial control of expression of the rhodopsins so that specific cell types can be examined. The current reliance on divisions between excitatory and inhibitory neurons (Gradinaru et al. 2009) or the use of non–specific markers (e.g. parvalbumin; Sohal et al. 2009) will not suffice.

4. DO WE HAVE TO GO BEYOND TRADITIONAL CRITERIA?

There are thus clearly significant challenges to face if we are to satisfy the minimal criteria for network understanding, challenges that currently cannot be dismissed by appealing to novel techniques. As systems increase in complexity, the number of specific components and functional properties that need to be examined becomes astronomical. Taking these networks apart is analogous to the ‘tyranny of numbers’ faced by electrical engineers in trying to assemble large numbers of individual components before the introduction of integrated circuits. Traditional and novel techniques all offer certain advantages to network analyses, but each works at specific levels and in limited domains, and linking between these is difficult. It is not simply that we need more data on networks, we also need new ways of thinking about and integrating the data we have to provide (if possible) general theories of network function. Overcoming the tyranny of numbers in electronics required integrative technologies. We may need analogous integrative concepts that allow us to recognize certain network motifs as obligatory functional units, thus helping us move beyond the need of identifying every component. These motifs could include network features such as feed-forward or feedback excitatory and inhibitory modules. However, this would require that the function could be reliably inferred when these features are present, and current insight gives little reason to be confident that this is the case (see Elson et al. 2002).

Many traditional problems could be removed by arguing that individual network components do not need to be identified, and that phenomenological principles of network operation (similar to mass-action effects) can be explained without recourse to lower-level properties (this would be analogous to statistical mechanics replacing Laplacian ideals of an absolutely deterministic classical physics; see Yates 1993). Focusing on average values of network, state variables and correlations between elements rather than individual components or their interactions can allow inferences of lower-level mechanisms or organization. However, assumptions have to be made about knowingly heterogeneous functional and structural properties. A wide range of potential mechanisms could be proposed depending on the initial assumptions used, all of which could be altered or negated by the influence of known or unknown elements or factors (Stevenson et al. 2008). These assumptions can only be constrained by direct insight into network properties and processing. The significant functional effects resulting from the activity in single neurons (Houweling & Brecht 2008) further suggests that the function cannot confidently be inferred without taking individual components into consideration. Nevertheless, it is currently impossible in anything but the simplest systems to come close to a complete characterization of all network components and their functional interactions. We thus have to rely on analytical approaches in order to understand integrative functions (discussed by Prinz (2010) and Smeal et al. (2010)), while at the same time needing greater correspondence with the underlying biological details.

While the traditional criteria for network understanding are hard to satisfy, they will certainly have to be extended. First, while it is obvious (if not Phil. Trans. R. Soc. B (2010)
always followed experimentally) that networks contain heterogeneous cellular and synaptic populations and effects cannot be extrapolated between them, it also has to be appreciated that variability also occurs within these populations (see Soltesz 2006). Variability was traditionally ignored in network analyses, population properties being reduced to average values. To some extent this was necessary: analyses are difficult and accepting that molecular, electrophysiological or anatomical markers only provide crude functional sub-divisions could make the situation seem hopeless. However, averaged values may not reflect any actual property of a cell or synapse (Golowasch et al. 2002), and collapsing all components within a population to an average value may turn out to be as simplistic as assuming that properties can be extrapolated between all classes of cells and synapses. Variability could reflect random variation in genetically identical cells at the level of transcription (Yu et al. 2006), plasticity-induced changes in control networks in different states, or programmed functionally relevant variability. The latter seems likely given that variability is a requisite of healthy physiological systems (Buchman 2002). Irrespective of the underlying cause, variability is a reality, and it means that what we assume of as defined populations based on anatomical features or molecular markers probably consist of functional sub-divisions that should have different influences on the processing of activity within a network (see Soltesz 2006). This raises a significant issue with respect to the identification of network neurons and associated experimental strategies. For example, in discussing the genetic dissection of neural circuits, Luo et al. (2008) define a cell class as a group of neurons that perform the same function. This is a fairly open definition that logically means that neurons traditionally defined as belonging to a single population based on shared anatomical or molecular properties should be considered different types of cell if their functional properties vary. However, it also means that a single neuron that made divergent connections with different properties onto different postsynaptic cells (a common feature; see Markram et al. 1998 and references therein) would belong to multiple functional classes depending on how its influence was measured (i.e. which postsynaptic cell was examined). Conversely, considering two neurons with the same function as members of the same class would not take into account degeneracy, a common feature of genetic, neural and evolutionary networks (Tononi et al. 1999). These, and other, aspects of variability complicate our classification of cell types, with obvious implications for the already difficult task of identifying network neurons (see above), and of using genetic or optogenetic tools. The number of functional sub-divisions in a population depends on what classification criteria are used and where classifying boundaries are drawn (see Parra et al. 1998). In addressing variability we need finer classifications than those offered by current anatomical or molecular markers to identify specific functional sub-divisions, and importantly we also need to understand the relevance of cellular and synaptic variability to network function. This will be impossible by monitoring, stimulating or knocking-out components while not distinguishing between distinct functional groups. As with averaging, this can only provide limited insight at best into the actual functional role.

A second largely neglected aspect in network analyses is glial cells. These were traditionally considered to have supporting roles (transmitter uptake, buffering extracellular K⁺ and H⁺) rather than active roles in the generation of network outputs. This view is seemingly no longer tenable: glial cells contain a range of voltage-dependent ion channels and receive synaptic inputs from neurons and in turn modulate synaptic transmission through direct electrical and chemical interactions. This has lead to the concept of the tripartite synapse (i.e. the presynaptic and postsynaptic neuron and their associated glial elements; Araque et al. 1999), which places glial cells as functional components of networks. As with neuronal variability, neuronal–glial interactions have largely been ignored in network analyses, but they must be addressed. This could introduce a major experimental load, not simply because many more components will be added, but because so much remains to be investigated in terms of glial subtypes, their intrinsic properties, their functional interactions with neurons and in particular their potential network roles. While glial effects have been studied at the level of single cells and synapses, they have not been considered to any great extent in a network context. An understanding of how these ubiquitous (and numerous) elements may influence network function is difficult to appreciate. Araque & Navarette (2010) will highlight this significant gap in our knowledge of network organization and function.

Neuronal variability and glial cells both extend the list of components that need to be addressed for network understanding. But in addition to adding components to the analysis, we may also have to consider our conceptual approach. Classical approaches to networks are reductionist and assume that the function of a network can be determined from knowledge of its component molecular, cellular and synaptic parts; that these parts exist in external relationships to each other; that they interact through deterministic chains of cause and effect; and that these interactions do not alter the basic component properties. The reductionist approach is actively debated in the philosophy of science, and in the physical sciences it has been known for many years that complex multi-component systems (i.e. neuronal networks) can exhibit macroscopic behaviours that cannot be understood even given complete microscopic knowledge (Anderson 1972). While cellular and synaptic variability, plasticity and glial cells add extra components to be analysed, they do not challenge the basic reductionist assumptions typically followed in network analyses. An obvious problem of the reductionist approach is that it is not constructivist: even if it were possible to reduce a network to its fundamental components it does not follow that you could then build up an understanding of global function (see above). Two aspects can briefly be considered that challenge reductionist views.

First, there is the long-standing question of whether networks are in principle deterministic (see Bullock...
1970). The answer could depend on the system, function, or conditions examined: some effects are highly predictable, while others are stochastic (probabilistic). Probabilistic effects could be introduced by molecular, cellular, synaptic or network noise, which is an inevitable component of biological systems (Smeal et al. 2010). This could be generated by voltage or ligand-gated channel noise or variability in transmitter release machinery. Noise can have significant functional relevance by possibly increasing the signal-to-noise ratio (stochastic resonance; Douglass et al. 1993) and enhancing weak or periodic inputs (Dorval & White 2005). It could also contribute to the neuronal variability discussed above. However, unlike deterministic variability resulting from definable state-dependent influences or functional sub-divisions of network components, probabilistic effects would prevent deterministic accounts of network function (Faisal et al. 2008). As Smeal et al. (2010) discuss, in addition to complicating experimental analyses and their interpretation, noise also sets limits on analytical strategies.

A second aspect is that while morphologically distinct network components clearly exist, they may be functional abstractions. Can the function of a neuron be determined if its analysis is divorced from the normal network structural and temporal aspects (e.g. a neuron is inhibitory or excitatory as a result of its effect on the postsynaptic neuron)? In thinking of network components as autonomous parts, we may lose sight of what nervous systems do and how we approach their analysis. We may need to move beyond thinking of neurons as building blocks that can be defined and understood in isolation to examining them on the basis of their specific spatial and temporal interactions. It is unlikely that anyone would deny the importance of these interactions, and it could be argued that this is a central point of network analyses, but in analysing these interactions, we still typically look for deterministic effects between autonomous components. While knowledge of system components and their interactions is essential for progress, it may still be insufficient to predict system function. This will be the case if the properties of the components are qualitatively altered by their interactions, thus generating emergent or self-organizing effects. This introduces the distinction between reducible systems in which computational shortcuts allow the behaviour of the system to be predicted, and emergent or irreducible systems in which the function cannot be formally calculated despite having well understood microscopic laws (see Binder 2009). These effects have been highlighted and discussed for many years in physical systems, which are more advanced in respect of being aware of the problems associated with reductionist analyses. It may be too soon for neurobiology to appreciate these problems and the limits that may eventually have to be faced. However, some appreciation of the potential for non-deterministic emergent effects may direct experimental and computational strategies without having to wait to hit a potential analytical ‘wall’.

The presence of emergent effects would necessitate a hermeneutic approach, where the whole is understood by reference to the parts and the parts by reference to the whole. Several aspects may demand this sort of analysis. First, there is the issue of ephaptic communication mediated by local chemical or electrical fields (Krnicewicz 1986; Jefferys 1995; Bokil et al. 2001; Kamermans & Fahrenfort 2004). These effects were assumed to have modest influences under physiological conditions (although significant roles were assumed in pathology), but activity-dependent changes in local electrical and ionic fields caused by changes in extracellular K⁺ and Ca²⁺ are known to significantly affect cellular and synaptic function (Erulkar & Weight 1977; Borst & Sakmann 1999; Bokil et al. 2001; Rusakov & Fine 2003; see Durand et al. (2010) for a discussion of these effects). These interactions will be influenced by the structural arrangement of network neurons and synapses as well as the tortuosity and anisotropy of the extracellular space (which represents approx. 20% of brain volume; Nicholson 2001). It can be claimed that these aspects can also be quantified in deterministic models by adding local field compartments (Barbour 2001). However, it seems unlikely that geometrical considerations can be divorced from functional and temporal aspects in this way, as network activity alters the volume of the extracellular space (by approx. 30%; Østby et al. 2009; Theodosis et al. 2008; see Araque & Navarette (2010) for the glial contribution to this effect). This introduces a circular causality: network and neuronal activity will be influenced by and will influence spatially dependent ephaptic effects.

Spatial, functional and temporal interactions could also directly influence transmitter effects. Network activity will evoke the release of various transmitters, resulting in an activity-dependent changing chemical field around network neurons. In addition to local simultaneous or co-release, transmitters can spill over to neighbouring synapses (Rusakov & Kullmann 1998) and could evoke volume effects by diffusing some distance from their point of release (Wood & Garthwaite 1994). The interactive effects of transmitters in this changing chemical field cannot be assumed from knowledge of their individual effects (see Brezina & Weiss 1997; discussed by Brezina 2010). In addition to their nonlinearity, these interactions will be influenced by the functional state of the releasing or target cells and synapses, as well as transmitter uptake, breakdown or diffusion mechanisms, factors that are influenced by and influence the properties of glial cells and their uptake mechanisms and the properties of the extracellular space (Genoud et al. 2006). This introduces the potential for numerous circular interactions: network activity will alter glial function and the extracellular space; this will in turn alter transmitter uptake and diffusion, thus altering glial function, the extracellular space, single and interactive transmitter effects and network activity. While circularity does not necessarily prevent understanding, these chains of circular interactions could lead to emergent effects that cannot be understood from knowledge of the individual components or their interactions under quiescent conditions (i.e. the traditional approaches to network analyses, approaches that the development of novel techniques seek to improve rather than revolutionize).
Electrophysiological, molecular and imaging techniques examine single components at a time and then attempt to integrate these serially acquired components into assumed temporal sequences. The problem is that there may be no sharp distinction between spatial, functional, and temporal aspects, and no one-to-one correspondence between lower-level and higher-level effects. The potential for this is again probably widely appreciated, but approaches still follow a mechanistic order that assumes understanding of the whole will follow from knowledge of the isolated component parts. We may need to invert this approach, and make the whole the fundamental from which parts emerge, that is, to appreciate that the components that we isolate analytically are at least to some extent created by the whole network. This is not simply to go through another cycle of switching from bottom-up to top-down approaches or to accept the view that effects can be understood without a knowledge of lower-level components, but instead to make the temporal and spatial order in which the components function and interact as important as the components themselves (i.e. a reciprocal structure–process relationship where structure and function are inseparable; see Bohm 1980). This goes beyond our currently assumed interactions, and would question whether network function can be understood if the normal spatial (electrical, ionic, and chemical fields) and temporal aspects are disturbed (e.g. if effects are examined in tissue slices, cultured or dissociated cells or in inactive networks, the routine and currently necessary approaches to network analyses). In this reciprocal relationship, no component would be autonomous, which will make it difficult to disentangle individual effects to provide deterministic accounts.

5. CONCLUSION

This brief introduction highlights traditional and novel approaches to neuronal network analyses. Measures of success ultimately depend on what we mean by understanding. This borders on the philosophical, but it is important for evaluating the success of the field and the claims that we can make, and of directing experimental strategies. The requirements seem intuitively simple: we need cellular-level resolution scaled up to multiple interacting cells so that we can characterize specific components in physiologically relevant spatial and temporal associations. This currently seems a distant goal given our current and emerging techniques. Overcoming the limitations either means resorting to logical fallacies where discussions are restricted to remove difficult or intractable analytical requirements: this is essentially what has been done in the systems that claim understanding (see above). What is included or excluded from network schemes or analyses should be openly debated, not assumed for convenience. Additionally, we need to develop experimental and conceptual insights that will bring difficult or intractable analyses within reach. In this case the merits of novel techniques should not be highlighted by playing down the aspects that it cannot address (e.g. functional analyses in connectomics), or by promoting competition between approaches that can both have utility (Scanziani & Hausser 2009).

Current analyses insist on significant trade-offs (e.g. precise functional analyses in cultured cells or tissue slices which require moving away from the normal structural, functional, or temporal environments). Analytical limits should not be dismissed by relying on assumptions of the sort encompassed by phrases like ‘everything else being equal’ or ‘in principle’, terms that have been used frequently in the past to justify specific approaches, techniques or claims. There is a temptation to draw a line under things, but we must recognize the aspects that we have not or cannot currently take into consideration. Given the influence of glia for example, can any network analysis that has not taken these into consideration be considered complete? And what else may be waiting to be discovered? As Ashby (1956) pointed out, a system can contain an almost infinite variety of variables, and any analysis by necessity picks out only those that are relevant to a particular study. Claiming something is characterized when a particular individual approach or interest has been satisfied becomes egotistical.

One significant problem, given that networks form the link between cellular properties and behaviour, is that analyses seldom examine networks in actual behaving systems. Even in intact preparations there are potential issues of restraint, recording devices or anaesthetics altering cellular and synaptic properties in ways that could complicate the identification of network neurons and synapses and their properties. In dissected preparations there is the issue of the extent to which fictive outputs correlate with behaviour or what behaviour they are trying to produce. This does not negate the utility of isolated network analyses; they provide insight into the network components of behaviour. But the relationship of individual network outputs to behaviour becomes increasingly unclear as systems increase in complexity and as experimental preparations become more reduced. In a similar vein, the stimuli used to examine networks become increasingly artificial as systems increase in complexity and become more reduced, and natural stimuli may evoke different effects to artificial inputs. Again, the aim is not to be pessimistic or nihilistic or to attack the utility of any approach or insight: we know that we are addressing complex analytical issues and even applying all the tools that we have network understanding is difficult. But this has to be addressed if we are to understand normal or pathologi
cal nervous system function and behaviour.

6. ‘MEDIOCRITY KNOWS NOTHING HIGHER THAN ITSELF’ (CONAN DOYLE)

While network reviews often focus on data and their interpretation, Yuste (2008) has recently gone further in addressing personal concerns about the sociological aspects of network analyses. This breaks with conventions that limit the discussion to conceptual and methodological issues (this would be sufficient if the science was completely rational and objective). However, Yuste has highlighted something seldom formally expressed but widely appreciated by those who practise and suffer from it, that is, that the peer review system allows dissenters or competitors with competing ideas to be blocked (if they get published, the last resort is to avoid citing them).
The converse also occurs, help is given to those seen as supporters, which has made networking a worryingly significant aspect of a scientific career. Because the analysis of neuronal networks is considered one of the major gaps in our understanding of the nervous system, there are potentially significant career and personal benefits to be won if the field can be convinced that a particular analysis or technique has (or will) lead to this understanding: this has probably fostered some of the exaggerated claims. Like Yuste, I have also seen people discouraged and leave the field as a result of aggressive reviewing and failure to cite work. I am not as optimistic as Yuste that changing publication practices will overcome the problem (the best that this could do is to address publication problems, failure to cite work and bias in grant review will remain). There probably is no solution beyond being aware of the potential biases we can all carry, and of being aware of the complexities and problems that we face so that caution is exercised over any significant claims.

I would like to thank Tom Gilbey, Erik Svensson and Jeremy Niven for discussions and helpful comments on the manuscript.

REFERENCES


Ding, J., Takasaki, K. & Sabatini, B. 2009 Suprareolution imaging in brain slices using stimulated-emission


Korn, H. & Faber, D. 1976 Vertebrate central nervous system: same neurons mediate both electrical and


constraints, uptake, and receptor activation. J. Neurosci. 18, 3158–3170.