Connecting connectomes

Maryann E. Martone[1] and **Giorgio A. Ascoli**[2]

Maryann E. Martone: maryann@ncmir.ucsd.edu; Giorgio A. Ascoli: ascoli@gmu.edu

[1]Department of Neuroscience, University of California, San Diego, CA, USA

[2]Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA, USA

New techniques and paradigms in neuroscience have drastically increased the amount and diversity of experimental measurements and simulation results to be handled and processed. This development requires new approaches for data storage, sharing, analysis, co-registration, search, and exchange. As often discussed in the pages of this journal, the design and implementation of broadly accepted and generalizable solutions will critically rely on both community-driven initiatives and standardization. The recent announcements of the Human Brain Project (HBP) by the European Community[1] and of the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) initiative by the US government[2] are expected to further accelerate the pace of data production in neuroscience and should thus exacerbate the need for neuroinformatics-based integration efforts. Both of these large-scale efforts, in fact, recognized the challenge from the start and promise to address it with somehow complementary strategies.

The European HBP (not to be confused with the homonymous US effort[3] that contributed substantially to the earlier establishment of the field of neuroinformatics) defines an ultimate goal that is computational at its essence: to create real-scale, neurobiologically realistic models of the mammalian brain in order to simulate its functions. The necessary assignment of numerical values to a very large number of parameters explicitly formalizes the underlying assumptions, at the same requiring an adequate neuroinformatics infrastructure. In fact, one of the main deliverables of the HBP will be its software infrastructure, thus emphasizing software development at least as much as computational modeling. The US BRAIN initiative, on the other hand, starts with a neurotechnology aim, to create as of yet lacking methods for comprehensive brain activity mapping. Again, it is immediately clear from the scale of the endeavor that ensuring data of such a magnitude are suitable for later use, even just at the level of appropriate annotation, will necessitate considerable parallel progress in neuroinformatics.

Both the HBP and BRAIN will devote a sizeable portion of effort and support to connectomics. Such emphasis is not surprising given the sheer extent of information richness in brain connectivity. As is becoming increasingly clear, neural circuitry remains highly complex at all scales of analysis from the macroscopic level of whole brain region-to-region connectivity to the microscopic level of neurons and synapses. Indeed, the term "connectome" from its original introduction[4] applied to multiple scales. In the last five years, two major sets of initiatives have been developed to tackle brain connectivity at the macroscopic and microscopic levels, respectively.

Macro-connectomic efforts involve both non-invasive brain imaging in humans and tractography in primates and rodents. The advantage of mapping the regional connectivity of whole brains in living human subjects is obvious and the scientific and social implications far-reaching. Leading techniques in this effort are diffusion[5] and functional[6] MRI. A major and ongoing boost to these efforts was provided by the NIH-funded Human Connectome Project, which supported two large research consortia[7]. These centralized developments are

paralleled by grass-root coordinated efforts such as the 1000 Functional Connectome projects[8] and a large number of distributed initiatives in hundreds of laboratories worldwide, many of which described in the pages of this journal.

A distinct set of large-scale, high-profile macroconnectomic initiatives involve whole-brain mapping of region-to-region connectivity in rodent models with tract-tracing, ranging from raw experimental data (including the Brain Architecture Project[9], the Mouse Connectome Project[10], and the Allen Brain Connectivity Atlas[11]) to knowledge bases[12].

At the microscopic level, several large-scale efforts are ongoing based on electron microscopy (EM), including the HHMI Fly EM Project[13], which aims to produce behaviorally-relevant synaptic-level circuits such as for the entire Drosophila nervous system, and the Ultrastructural Brain Mapping Consortium, which targets the whole mouse brain[14] starting with selected circuits[15]. EM allows the dense reconstruction of each and every synapse in the circuit. Conceptually similar levels of analysis can be carried out with light microscopy, especially using molecular visualization techniques, such as with super-resolution[16], array tomography[17], GRASP[18], viral transfection[19] or genetic bar-coding[20].

Despite the great methodological diversity at both ends of the scale (from fMRI to tractography and from EM to confocal microscopy), most ongoing large-scale connectomic efforts can be thus grouped into two broad approaches based on the conceptual target of investigation: at the macroscale level, the nodes of the connectivity graph are brain regions (functional areas, voxels, cortical columns, etc.), and the edges are projections among them. At the microscale level, the nodes are neurons and the edges are synapses or close axon-dendrite proximities (also known as potential synapses[21]). To avoid the ambiguity of the word "connectomes" as referring to both of these types of circuits, the terms "projectomes[22]" and "synaptomes[23]", respectively, have been proposed. Perusal of the last few volumes of this journal suggests that connectomics may be a step or two ahead of synaptomics with respect to standardization effort. This may potentially cause problems down the line massive amounts of synaptomics data are broadly expected in the not too distant future.

Today, the projectomic and synaptomic research worlds are practically separate. This is reasonable, given the complementary domains of investigations, limits, and benefits of these endeavors. Projectomics allows for non-invasive whole-brain mapping in humans, but its explanatory power is limited. Synaptomics gets to the core mechanism of neural computation, but its range of applicability in the foreseeable future is limited to simple model organisms or minute fractions of mammalian nervous systems[24]. In many ways, synaptomics starts where projectomics leaves off (and vice versa), as the entire volumetric reach of ongoing synaptomic efforts (1 mm^3) corresponds to a single unitary element in a projectome.

While understandable, the current disconnect between projectomics and synaptomics is nonetheless less than ideal. The two levels of connectomics share one and the same quantitative framework (graph theory) as well as several notably common themes, such as properties of small world network, non-uniform distributions, and relevance of hubs. Most importantly, projectomes and synaptomes are two representations of the same objects, namely nervous systems. Many if not most foundational scientific discoveries in all disciplines came from linking phenomena at different scales, such as classical mechanics and thermodynamics, electromagnetism and optics, quantum theory and chemistry, genetics and genomics. It is tempting to speculate that the unification of projectomics and synaptomics will be a crucial step in the path to understand the brain and its relationship to behavior and the mind[25].

We purport that neuron types may be the suitable conceptual bridge to connect projection- and synaptic-connectomes. The basic distinction between projection neurons and local interneurons underscores the link between these two levels of analysis. Local interneurons are only relevant to synaptomes and their domain is entirely confined within a single projectome voxel. Projection neurons, in contrast, provide the cellular substrate for macroconnectomes. If two regions are connected in the projectome, there must be at least one type of projection neuron whose axonal arbor corresponds to the reported tract. A neuron-type connectome thus provides the useful mesoscale between synaptomes and projectomes.

The central questions of neural cell types, what defines them, how many there are, how to model them, still constitute largely open issues in the neuroscience community. Without a reasonable definition of neural cell types, we lack the cellular building blocks for establishing tractable models that integrate micro- and macro-circuitry. At the same time, neuron types constitute a core tenet of both experimental and computational neuroscience, and are pivotal to quantify neural diversity, developmental and pathological changes, inter-subject variability, and species specificity. The notion of neuron types is intrinsically statistical, and as such naturally amenable to empirical sampling and gradual transition to high-throughput data acquisition. One of the critical goals of neuroinformatics is to provide information technology infrastructures for dealing with big data, which is essential to bringing scales together. Thus, it is easy to predict the need for neuroinformatics in leveraging neuron types to connect connectomes. Relevant resources are already in place, such as the Neuroscience Information Framework[26] (NIF) and Neurolex.org[27], which are well positioned to foster a much required community involvement for the eventual success of such integration.

In recognition of the importance of formal models for defining neuronal cell types, the International Neuroinformatics Coordinating Facility[28] (INCF)targeted the representation of neurons as one of the core activities of the Program in Ontologies on Neural Structures[29]. The Neuron Registry Task Force assembled an international panel of experts in a series of meetings to discuss what constitutes a neuron cell type[30] and experimented with various interfaces and strategies for developing a knowledge base of neurons based on their properties[31]. These attempts suggested that an effective first step would be to identify a set of well-defined neuron types within key brain regions. To this aim, the ongoing Neurolex Neuron effort focused on an initial group of properties that could be annotated across several hundred neuron types. All of these neuron type properties are being entered into the Neurolex wiki, so as to immediately become publicly accessible and integrated within the knowledge base. Because Neurolex is built on a semantic wiki, each neuron is linked via its properties to other category pages, e.g., brain regions, molecules and morphological properties. Each neuron type is given a unique identifier so that it can be unambiguously recognized within publications and other databases or knowledge bases, such as of morphological reconstructions[32] or of physiological properties[33].

Some fundamental problems in neuroscience require a collective effort to solve. Thus, it is important to use a public space and a community effort to engage the neuroscience community in discussions about connectivity across scales, techniques, and species. As evident from initial efforts, this task is extremely difficult, as it forces researchers to confront the assumptions about key parameters from each of the sub-disciplines comprising neuroscience. Thus, the physiologist, molecular biologist, modeler, anatomist, developmental biologist, all have a slightly different perspective on the important parameters from their domains that are critical for defining a cell type and a circuit. Without the ability to expose these different perspectives, we are at risk of continuing to create more and more detailed, but largely disconnected, descriptions.

The most critical challenge in engaging the neuroscience community to publicly define cell types is the requirement to limit or even avoid the use of free text to ensure machine readability of the resulting product. Forcing researchers to confront the details of their conceptual universe in rigorously logical terms sheds light on the strengths and needs of the current models of anatomy, physiology, and molecular biology. The Semantic Wiki is a suitable platform for such an interaction, as it allows changes by multiple users on entities within their domain of expertise, propagating this knowledge throughout the system and highlighting disconnected statement and entities, which point to missing pieces of required knowledge.

Web search engines like Google and social networks like Facebook learn a lot from analyzing the behavior of their users, and not just the content of web pages and profiles. Likewise, wikis enable not only mining of the information provided, but also of the behavior of the contributing experts, which can foster a better understanding of the current state of knowledge in the neurosciences. For example, the wiki automatically tracks changes and allows users to compare different versions. Anyone can thus extract fields or entries with heavy editing or comments posted, suggesting that some confusions or disagreements may exist for that particular entity. In other words, tracking the history of changes may indicate areas of disagreement or model shortcoming. The goal of such a knowledge space would not be to provide certainly where uncertainty exists, but to create a public place for exposing this uncertainty to machine-based approaches.

Although there is clearly a learning curve for wiki-editing, the benefits of having everyone work within a web accessible platform, without a filter between the expert and the informatics system, are significant. Each entry in the lexicon represents a model of our current thinking and uses the concepts bandied about by neuroscientist every day to express how we think that the nervous system is organized. The challenge is to link these models to active data collection to test, refine or discard them. The knowledge space becomes a living, computable view of current thinking and an index into the myriad forms of data that bear upon the asserted models. This task is beyond the ability of any one individual, group or institution. Moreover, unlike traditional databases, the Wiki platform makes this knowledge and data searchable via major search engines. This ensures that these pages will be found by neuroscientists searching the web, and by non-experts searching for neuroscience information, without them needing to visit a particular website.

A comprehensive, dynamic, online description of neuron types can thus reveal not only how they relate to synaptic microcircuits and macro-connectomes (e.g. do our current conceptualizations of connectivity among neurons account for the levels of organization observed in fMRI?), but also the interaction among groups of researchers (e.g. what questions or data do computational modelers need from the experimental community and vice versa?). This level of integration requires a community effort to define an information model to relate resting state fMRI to tractography, projections from intracellular tracers to single neuron studies, both functional and structural. A knowledge space where these discussions can take place and the data exposed is also an ideal driver to provide training opportunities in neuroinformatics, coordinate community-driven efforts for creating standards and procedures for their solution, promote interoperability of existing digital resources, and further data and resource sharing.

To date, over 150 individuals have contributed their work to the Neurolex, which has received collectively over 200,000 edits. In particular, the ~250 neuron types entered so far[34] involved contributions from more than 40 neuroscientists[35]. This seed can be expanded by progressively broader involvement of the community through workshops structured as "knowledge hackathons", in which attendees work within the knowledge space to learn how

to link their data to the evolving conceptual structure and how to modify such structure itself in light of discussions. Initially, both neuroinformatics experts and neuroscientists will likely be involved in defining the requirements for adding information and data to the knowledge space.

The power of neuroinformatics tools and open science is that they are accessible to everyone with access to a computer. The advent of the internet created an unprecedented opportunity to move science to a more open, cooperative platform that allows more effective leverage of data, tools and expertise. The time has come to cease the endless specialization within narrow domains of neuroscience, where we accrue ever finer details about the same types of entities, collected in information silos like papers and individual databases. To move neuroscience forward, it is time to move neuroscience into the web where knowledge and data can be shared and linked in a machine-accessible way.

References

1. humanbrainproject.eu

2. en.wikipedia.org/wiki/BRAIN_Initiative

3. De Schutter E, Ascoli GA, Kennedy DN. On the future of the human brain project. Neuroinformatics. 2006; 4(2):129–30. [PubMed: 16845164]

4. Sporns O, Tononi G, Kötter R. The human connectome: A structural description of the human brain. PLoS Comput Biol. 2005 Sep.1(4):e42. [PubMed: 16201007]

5. Jones DK, Knösche TR, Turner R. White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. Neuroimage. 2013 Jun.73:239–54. [PubMed: 22846632]

6. Breakspear M. "Dynamic" connectivity in neural systems: theoretical and empirical considerations. Neuroinformatics. 2004; 2(2):205–26. [PubMed: 15319517]

7. humanconnectome.org and humanconnectomeproject.org

8. fcon_1000.projects.nitrc.org – see also Kennedy DN. Making connections in the connectome era. Neuroinformatics. 2010 Jun; 8(2):61–2. [PubMed: 20428970]

9. brainarchitecture.org

10. mouseconnectome.org

11. connectivity.brain-map.org

12. braincusi.org – see also Bota M, Dong HW, Swanson LW. Brain architecture management system. Neuroinformatics. 2005; 3(1):15–48. [PubMed: 15897615]

13. janelia.org/team-project/fly-em

14. connectomes.org

15. neuro.mpg.de/connectomics

16. Maglione M, Sigrist SJ. Seeing the forest tree by tree: super-resolution light microscopy meets the neurosciences. Nat Neurosci. 2013 Jul; 16(7):790–7. [PubMed: 23799471]

17. Micheva KD, Bruchez MP. The gain in brain: novel imaging techniques and multiplexed proteomic imaging of brain tissue ultrastructure. Curr Opin Neurobiol. 2012 Feb; 22(1):94–100. [PubMed: 21944260]

18. Kim J, Zhao T, Petralia RS, Yu Y, Peng H, Myers E, Magee JC. mGRASP enables mapping mammalian synaptic connectivity with light microscopy. Nat Methods. 2011 Dec 4; 9(1):96–102. Nat Methods. 2011 Dec 4;9(1):96-102. [PubMed: 22138823]

19. Callaway EM. Transneuronal circuit tracing with neurotropic viruses. Curr Opin Neurobiol. 2008 Dec; 18(6):617–23. [PubMed: 19349161]

20. Zador AM, Dubnau J, Oyibo HK, Zhan H, Cao G, Peikon ID. Sequencing the connectome. PLoS Biol. 2012; 10(10):e1001411.10.1371/journal.pbio.1001411 [PubMed: 23109909]

21. Stepanyants A, Chklovskii DB. Neurogeometry and potential synaptic connectivity. Trends Neurosci. 2005 Jul; 28(7):387–94. [PubMed: 15935485]

22. Kasthuri N, Lichtman JW. The rise of the 'projectome'. Nat Methods. 2007 Apr; 4(4):307–8. [PubMed: 17396125]

23. DeFelipe J. From the connectome to the synaptome: an epic love story. Science. 2010 Nov 26; 330(6008):1198–201. [PubMed: 21109663]

24. See e.g. openconnectomeproject.org

25. Ascoli GA. The Mind-Brain Relationship as a Mathematical Problem. ISRN Neuroscience. 2013; 2013:13 pages. Article ID 261364. http://dx.doi.org/10.1155/2013/261364.

26. neuinfo.org

27. Larson SD, Martone ME. NeuroLex.org: an online framework for neuroscience knowledge. Front Neuroinform. 2013 Aug 30.7:18. [PubMed: 24009581]

28. incf.org

29. pons.incf.org

30. incf.org/programs/pons/neuron-registry

31. Hamilton DJ, Shepherd GM, Martone ME, Ascoli GA. An ontological approach to describing neurons and their relationships. Front Neuroinform. 2012; 6:15. [PubMed: 22557965]

32. NeuroMorpho.org

33. neuroelectro.org

34. neurolex.org/wiki/Category:Neuron

35. neurolex.org/wiki/Neuron_Curation_Home