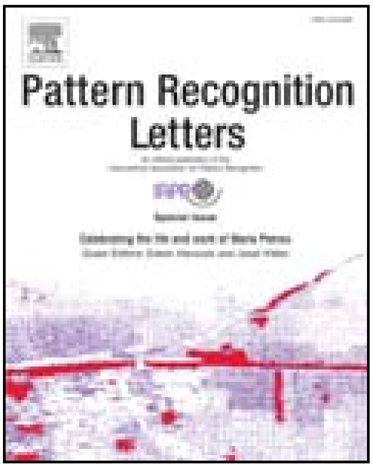
Accepted Manuscript



On Unifiers, Diversifiers, and the Nature of Pattern Recognition

Gavin Brown

PII: S0167-8655(15)00131-2
DOI: 10.1016/j.patrec.2015.04.014
Reference: PATREC 6213

To appear in: *Pattern Recognition Letters*

Received date: 18 July 2014
Accepted date: 18 April 2015

Please cite this article as: Gavin Brown, On Unifiers, Diversifiers, and the Nature of Pattern Recognition, *Pattern Recognition Letters* (2015), doi: 10.1016/j.patrec.2015.04.014

On Unifiers, Diversifiers, and the Nature of Pattern Recognition

Gavin Brown

School of Computer Science, University of Manchester, UK

Abstract

We study a dichotomy of scientific styles, *unifying* and *diversifying*, as proposed by Freeman J. Dyson. We discuss the extent to which the dichotomy transfers from the natural sciences (where Dyson proposed it) to the field of *Pattern Recognition*. To address this we must firstly ask what it means to be a "unifier" or "diversifier" in a field, and what are the relative merits of each style of thinking. Secondly, given that Dyson applied this to the *sciences*, does it also apply in a field known to be a blend of science and engineering? Parallels are drawn to Platonic/Aristotelian views, and to Cartesian/Baconian science, and questions are asked on what drives the Kuhnian paradigm shifts of our field. This article is intended not to marginalise individuals into categories (unifier/diversifier) but instead to demonstrate the utility of philosophical reflection on our field, showing the depth and complexities a seemingly simple idea can unearth.

1. Introduction

In his 1988 book, *Infinite in All Directions*, the theoretical physicist Freeman J. Dyson discusses two distinct styles of scientific thinking: *unifying*, and *diversifying*, claiming that most sciences are dominated by one or the other in various periods of their history.

> "**Unifiers** *are people whose driving passion is to find general principles which will explain everything. They are happy if they can leave the universe looking a little simpler than they found it."*
> *"***Diversifiers** *are people whose passion is to explore details. They are in love with the heterogeneity of nature* [...] *They are happy if they leave the universe a little more complicated than they found it."* (Dyson, 1988, ch. 3, pg. 44)

When I first read these quotes, and showed them to colleagues, there were a number of immediate assumptions. For example, some assumed that unifiers are *theoreticians*, and the diversifiers are *experimenters*. Others took the dichotomy to be equivalent to *scientists* vs *engineers*, or to *academia* vs *industry*. In association with intellectual endeavour, the terms *unifying* and *diversifying* seem to come with a certain semantic "baggage". This is exemplified by the media-fuelled furore surrounding 20th century physics, with science celebrities seemingly promoting a unifier viewpoint and the search for the 'ultimate laws of the universe'.

Dyson's treatment of this is relatively short, at just one 18-page chapter (Dyson, 1988). It is therefore important to moderate our contemporary biases, if we are to understand what he intended. It seems appropriate to engage in a conceptual analysis of these terms, with a major question being whether they have the same meaning in *natural* sciences (where Dyson conceived them) as they do in a *computational* science like our own.

Whilst Dyson expands upon his view with examples from physics and biology spanning 400 years, our own field of Pattern Recognition[1] is relatively young. If

Email address: `gavin.brown@manchester.ac.uk` (Gavin Brown)

[1] In this article I make no distinction between the field of *Machine Learning* and that of *Pattern Recognition*, as this has been addressed elsewhere. I choose the term PR simply because of the name of this journal.

physics is the old man of science, then we're the spotty teenagers. This considered, it is good to look back and consider how far we've come, where we're going, and whether we can learn something from the older disciplines. This type of philosophical reflection allows us to plan objectives, to understand our motivations, successes and failures, both as a collective and in our individual pursuits. The purpose of this article is to reflect in this way, on how the dichotomy transfers from the natural sciences to a science of *computation*, and more specifically, *Pattern Recognition*.

1.1. What are we?

Pattern Recognition is a multifarious field. We study the science and engineering elements of *data*. We are interested in automating the understanding of data, including *prediction* and *description* of phenomena. The construction of both heuristic and formal mathematical models forms the backbone of our culture. The field was spawned from the dreams of Artificial Intelligence, though the reality has encompassed a far broader scope of study than originally envisioned at the Dartmouth Conference (McCarthy et al., 1955). However, we are not tackling the wider integrative challenge of A.I., but instead focused on a restricted (yet immensely challenging) problem: the automated processing and inference problems that arise from diverse sources of data. At present, we encompass aspects of pure/applied statistics and mathematics, computer science, and biologically-inspired mechanisms, among others.

1.2. Structure of this Article

As mentioned, Dyson's terminology of 'unifiers' and 'diversifiers' lends itself to a number of potential implicit meanings – a deeper analysis of these is a necessary first step, tackled in Sections 2 and 3 of this article. Section 4 will explore how the dichotomy transfers over to Pattern Recognition. For example, something quite explicit from Dyson's writings is that he equates *unifying* with *simplicity*. This reflects his training in physics, where the belief is widespread that beautiful (or simpler) theories are more likely to be correct. But what does this mean in Pattern Recognition, and how is it different than in natural sciences like physics or chemistry? Sections 5, 6 and 7 consider the nature of work in our field, and of how revolutions in a field come about – are they driven by unifiers, or diversifiers, or both?

Finally, section 8 will play Devil's Advocate, and ask *why study this?* What is the *value* of the dichotomy as a conceptual tool? What benefits may come, to the individual or to the community, from addressing these philosophical questions?

2. Two Styles of Thinking

Dyson states that *unifiers* (citing Albert Einstein as the exemplar) believe the universe can be reduced to a finite set of principles — a simple, elegant framework, couched in the language of mathematics — and have the pursuit of this as their primary goal in science. On the other hand, *diversifiers* (citing Emil Wiechert, a geophysicist who discovered the layered structure of the Earth) prefer to explore the infinite diversity of details in the universe, often creating new phenomena and tools simply for the sake of exploring those details. Wiechert delivered a lecture in 1896 in which he stated:

> *"So far as modern science is concerned, we have to abandon completely the idea that by going into the realm of the small we shall reach the ultimate foundations of the universe. I believe we can abandon this idea without any regret. The universe is infinite in all directions, not only above us in the large but also below us in the small. If we start from our human scale of existence and explore the content of the universe further and further, we finally arrive, both in the large and in the small, at misty distances where first our senses and then even our concepts fail us."*

Einstein, as a unifier, believed the large and small of the universe could be *abstracted* into a single unified theory. Wiechert, as a diversifier, believed the universe is inexhaustible and potentially incomprehensible to the human mind — that no matter how long or far we look into the "misty distances", the universe *will not conform to abstractions.* For a diversifier, the details matter more than the simple explanations. Where a unifier prefers abstract structure and the aesthetics of a unified theory, the diversifier focuses on the concrete variations of nature, the exceptions to the theory.

This dichotomy could be (mis-)interpreted in several ways. One could read it as equivalent to theoretician/experimenter, to academia/industry, or to scientist/engineer. Or, taking in a broader philosophical context: to Platonic/Aristotelian views, Cartesian/Baconian science, or analysis vs synthesis as processes for generating knowledge. The following sections will argue that none of these is exactly isomorphic to Dyson's dichotomy; but, on deeper reflection, all provide fascinating perspectives for our own field.

2.1. Theoreticians and Experimenters?

On a first reading, it could be perceived that Dyson's unifiers are *theoreticians*, while diversifiers are *experimenters*. This is reinforced by his naming of the great experimenter Ernest Rutherford as a diversifier. Rutherford was an outstanding experimental physicist, but according to Dyson, disrespectful of academic learning, more interested in facts than theories. Rutherford was well known for statements such as *"If your experiment needs statistics, you ought to have done a better experiment"*, and referring to theoretical physicists he once joked *"they play games with their symbols, but we turn out the real solid facts of nature"*. Rutherford's diversifier perspective on scientific research provided new capabilities, such as determining the size of an individual atom, or counting the number of atoms in a given volume of gas. To be clear: his purpose was *science*, not engineering, but he was interested in the "real solid facts of nature". Dyson states that Einstein and Rutherford held such opposing views, greater than the normal rift between theorist/experimenter, that they could barely talk to each other — explained by the fact that they held fundamentally different philosophies on the nature and purpose of science.

However, assuming Einstein/Rutherford as the definitive unifier/diversifier split does not appear to be Dyson's intention. The simple mapping of *unify = theory* and *diversify = experiment* is far too naïve. He explicitly names a theoretical physicist, John Wheeler, as a diversifier. Wheeler (1911-2008) was one of the most prolific and accomplished theoretical physicists of the 20th century, a pioneer in quantum gravity and the theory of nuclear fission, he also introduced the term 'black hole', and 'wormhole' to describe hypothetical tunnels through space-time. He was also an early advocate of the "anthropic principle" – that the laws of physics are fine-tuned for the existence of life in the universe. However, Wheeler suggested a stronger extension, the *participatory* anthropic principle, in which the laws of physics are not primary, but derivative, and *brought into being* by the presence of conscious life in the universe. Here, Einstein's pure reductionist approach to physics, hunting for a single unifying set of laws, is turned on its head — the laws themselves are *mutable*, and are a function of our observation. Thus the search for unifying laws may be futile, since we cannot observe other laws that may have come into existence without us. Of course, it could be that Wheeler saw a deeper set of *developmental* governing laws. But, the very fact that the rest of the physics community was converging on a single unifying theory, and Wheeler challenged their viewpoint by bringing into the equations the 'tiny' detail of their own consciousness, makes him a *diversifier*. In Dyson's words:

> *"Among contemporary physicists, John Wheeler is unique in taking seriously the possibility that the laws of physics may be contingent upon the presence of life in the Universe."* [...]
> *"Wheeler's colleagues love him more than they listen to him. The physics of the unifiers has no room for his subversive thoughts."*

So, now we have a theoretician-diversifier. It is also easy to think of the converse, a unifier who relies on experimental observation. Charles Darwin's approach was almost exclusively observational and empirical in nature; in his autobiography he reflects on his career as so:

> *"Therefore, my success as a man of science, whatever this may have amounted to, has been determined, as far as I can judge, by complex and diversified mental qualities and conditions. Of these the most important have been* [...] *industry in observing and collecting facts, and a fair share of invention as well as of common-sense."*
> (Darwin, 1887, p144)

It is widely acknowledged that by the word 'invention', Darwin meant invention of hypotheses that can be experimentally tested. Darwin held a unifier mindset, reducing our very existence to the result of a single principle (natural selection) yet every step of the work relied on observation and experiment.

It is equally easy to name an experimental *physicist* with a unifier mindset. Though Dyson did not explicitly name him, one posits that James Prescott Joule (1818-1889) would be a typical unifier in his mind. Joule determined equivalencies between thermal, electrical, and mechanical phenomena, through rigorous experimentation. His principle of energy conservation is not a product of esoteric mathematics, but careful control of external factors that could have affected his experimental observations. This work unified numerous competing viewpoints, laying the foundation for the modern theory of thermodynamics. So, we have a theoretician-diversifier, and an experimenter-unifier, and the converse for each case.

	Theoretician	**Experimenter**
Unifier	Einstein	Darwin / Joule
Diversifier	Wheeler	Rutherford

While Dyson makes clear that instances of theorist-diversifiers are possible, it seems to be his contention that in his own field they are rare, and 20th century theoretical physicists are more likely to be unifiers. Equivalently, though we have two examples (Darwin/Joule), one posits that experimenter-unifiers (in any field) are rare. Thus, it may be that *most* unifiers are theory-oriented, but not all, and *most* diversifiers are more experimental, but again, not all.

2.2. Academia and Industry?

To underline the nature of his unifier/diversifier dichotomy, Dyson presents an analogy, rephrasing it in social terms.

> *"The first academic city in the world was Athens, and the first industrial city was Manchester, so I like to use the names of Athens and Manchester as symbols of the two styles of scientific thinking."* (Dyson, 1988, p37)

He clarifies later,

> *"The science of Athens emphasises ideas and theories; it tries to find unifying concepts which tie the universe together. The science of Manchester emphasises facts and things; it tries to explore and extend our knowledge of nature's diversity."* (Dyson, 1988, p40).

To clarify, he is not stating that all academics are unifiers, nor that all of industry are diversifiers. Neither is he explicitly stating that Manchester's *industry* was the home of the diversifiers he refers to. In fact he refers equally to the practice of science no matter where it occurs, within academic walls or in industry.

> *"Science belongs to both worlds, but the style of academic science is different from the style of industrial science. The science of the academic world tends to be dominated by unifiers, while the science of the industrial world tends to be dominated by diversifiers."* (Dyson, 1988, p36).

The qualification "tends to be" is important here. Whist he says industry "tends to be dominated" by diversifiers, it is interesting to consider the cause of this — whether diversifier-style science is *a function of industrial requirements*. The Manchester exemplar is particularly illuminating in this respect, given a deeper look at its historical context. Manchester, situated in the North of England, was the birthplace of the Industrial Revolution, and the growth of its intellectual capital is well documented by Thackray (1974). In the late 18th century, a number of learned societies[2] were founded by a group of dissatisfied intellectuals driven by a common vision – to escape the constraints forced upon the North from the wealthy elite in the South of England, typified by men with classical Oxford and Cambridge educations. While the industrial revolution was an obvious driver of science in this period, Thackray argues that the immense scientific innovation of the age was as much a means for

> *"the social legitimation of marginal men,* [...] *the adoption of science as a means of cultural expression by a new social class"* (Thackray, 1974, p678).

Dyson summarises neatly that "*the atmosphere of Manchester was saturated with contempt for the ancient universities"*. The style of science in these learned societies (and ultimately the University) was mirrored by these geo-social pressures.

> *"Science did flourish in Manchester during the crucial formative years of the industrial revolution, but*

[2]The first among which was the Manchester Literary and Philosophical Society (1781), which ultimately lead to the foundation of Owens College (1851), later renamed The University of Manchester.

> [...] *did not arise in response to the needs of industrial production. The driving force of the Manchester scientific renaissance were not technological and utilitarian; they were cultural and aesthetic."* (Dyson, 1988, p38)

He argues that, although Manchester's industrial needs were evident, the diversifier scientific style was not a strict function of this need. Instead it was borne out of a need for cultural distinction from the traditional Universities; and, in the wider city "to raise the aspirations of leadings citizens to a loftier level", showing that it was possible to live in Manchester and still be a gentleman. One element of Dyson's diversifier seems therefore to be a "rebellious" nature, to sit in opposition to convention, finding exceptions where others assume none.

Dyson names his archetypal diversifier of 19th century Manchester as Ernest Rutherford, whose academic work in understanding the structure of the atom was wholly curiosity-driven, without thought of immediate applications. Rutherford was well known to challenge convention, to ignore elegant theories in favour of observable facts. This combination of traits seemed to work for him, providing strong foundations for the emerging field of thermodynamics. This of course had great implications for the industrial revolution, though the industrial side was in full swing before Rutherford came along.

One notable omission from Dyson's argument is the distinction between *science* and *engineering*. Given the discussion on Manchester being the first industrial city, with many aspects of its activity oriented toward engineering, this is surprising. This is also especially relevant if we are to see whether this applies to Pattern Recognition, commonly viewed as a field straddling both science *and* engineering.

2.3. Science and Engineering?

It is here we must be careful with our dichotomies. Academia and Industry are *venues*, not practices. Science and Engineering are *practices*. Unifying and diversifying are *styles of practice*. These, according to Dyson, occur in both academia and industry in the practice of science, and as will be argued here, also the practice of engineering.

Dyson made no explicit statements on whether his dichotomy was intended to apply only to science, but we can speculate. The debate over the definition of *science* versus *engineering* could form an entire article by itself, and an excellent example is to be found within this special issue (Pelillo et al., 2015). Acknowledging this, but for the purposes of simplicity, I will adopt a distinction as so:

science is a practice primarily concerned with *truth*,

engineering is a practice primarily concerned with *utility*.

A naïve step would be to assume unifiers = scientists; however, Dyson's concept of a unifier does not seem to preclude the possibility of them having an engineering mindset. A unifier is someone who thinks about the relations *between* artefacts, rather than artefacts themselves. A unifier is, like Einstein and Darwin, concerned with how much of the universe can be brought under their metaphorical umbrella. One can imagine an engineer taking on the challenge of building a bridge, but with a unifier's view. This engineer would be concerned with characteristics of bridges that make them all strong, with the physics of how they can be modelled in a variety of situations, as opposed to the nuances of how one particular bridge should be built. In our own field of Computer Science, a unifier-engineer may be concerned with building *frameworks* of software/hardware or mathematics, for others to use, incorporating as many general principles as possible. This may sacrifice functionality in favour of having a clean single interface to a number of underlying tools. A diversifier-engineer in our field is more concerned with pushing limits, testing when and where individual techniques do or do not work — for example, they may demonstrate how parameter settings can be found automatically using efficient mechanisms imported from other fields, or evaluating scenarios where the mode of application for a predictive model is not so clear-cut.

In conclusion, Dyson's dichotomy easily applies across the science/engineering boundary. Whether one is interested in the pursuit of pure scientific knowledge, or of more practical goals, this does not limit a person to one *style* of research thinking.

2.4. Summary

Unpicking Dyson's dichotomy, a unifier is someone who is comfortable making abstractions or assumptions in order to reach a broader conclusion; where unifiers tolerate abstractions, diversifiers question them, and pursue the details; unifiers emphasise *similarities*, while diversifiers emphasise *differences*. Both these styles can be followed in academia or in industry, by theorist or experimentalist, by scientist or engineer. Whilst mathematics is a strong element of a unifier's toolbox, it is not *the* defining element.

The unifier's assumptions may be questionable, but the reasoning process followed from them is not. This approach allows great leaps of thought, by abstracting away from potentially flawed observations to an idealised form. Diversifiers on the other hand, cannot ignore the concrete variations of nature. To them, compromise or conformity to dogma seems alien, ignoring the observable facts as they can plainly be seen. They love the details, they see and enable things unifiers cannot, simply by persistence, fertile imagination, and systematic thought.

Given this breakdown of the concepts, some parallels to established dichotomies in classical and modern philosophy can be seen. In particular the idea of unquestionable reasoning from base assumptions is effectively the deductive process, championed by Descartes, and the idea of abstracting away details to have an "idealised form" is reminiscent of Plato's worldview. In the following section we will discuss these parallels.

3. Parallels to Philosophical Literature

3.1. Plato's Forms versus Aristotle's Empiricism?

Dyson states that the science of Athens (unifiers) emphasises "ideas and theories", whereas the science of Manchester (diversifiers) emphasizes "facts and things". The most immediate philosophical parallel here is Plato versus Aristotle.

At his most fundamental, Plato's position was that, progress toward new knowledge only begins when we come to think of our world experiences as flawed and possibly irrelevant, and it is only by processes of abstract thinking that we generate true understanding. Plato believed that humans were superior, born with *innate* knowledge, from which the full truth and knowledge of the universe (including theories of society, justice and government) could be reached by deduction alone, thus reference to empirical data was unnecessary, even distracting.

Aristotle on the other hand believed inductive processes could be used to establish first principles, combined with abstractions only when justified, from which deduction could be trusted and results later tested. This fits better with modern scientists arguably of the unifier mindset — Hawking, Feynman, and Dyson himself, proposed theories that can be tested, built on the foundations of observable phenomena. Feynman once delivered a memorable speech on the meaning of the modern scientific method:

> *"If it disagrees with experiment, it's WRONG. In that simple statement is the key to science. It doesn't make any difference how beautiful your guess is, it doesn't matter how smart you are, who made the guess, or what his name is. If it disagrees with experiment, it's wrong. That's all there is to it."* (Feynman, 1964)

Whilst Feynman's Nobel prize-winning work (Quantum Electrodynamics) undoubtedly involved a "unifier" perspective on physics, he did not require the level of abstraction in thought that Plato would have insisted upon. On the other hand, Aristotle's writing on science (natural philosophy) was wholly qualitative, he simply did not have access to quantitative tools like clocks or thermometers to *measure* the universe. As a consequence Aristotle was observational but not strictly in the sense of modern scientific method. Thus, in his own temporal context, Aristotle was probably a diversifier, but the details that would be scrutinised by a modern diversifier were perhaps unintentionally glossed over by the Aristotelean worldview.

3.2. Cartesians and Baconians?

Almost at the outset of his essay, Dyson seems to equate his dichotomy with Cartesian versus Baconian science:

> *"Historians of science are accustomed to call these two traditions in science Cartesian and Baconian, since Descartes was the great unifier and Bacon the great diversifier at the birth of modern science in the seventeenth century."* (Dyson, 1988, p40)

However it is doubtful that Dyson believes these are exact synonyms for his terminology. In later a communication he states that only *"roughly speaking, unifiers are following the tradition of Descartes, diversifiers are following the tradition of Bacon"* (Dyson, 1989). The "rough" correspondence between unifier/diversifier and Cartesian/Baconian is supported by further unpicking of the concepts.

Descartes followed the *rationalist* view, that the universe has an inherently logical structure, and its entirety could be deduced from first principles. The belief in the strength of the deductive process, and a belief that there exists an underlying logical structure to pursue, are strong unifier traits. However, the defining tenet of Descartes' philosophy was his *Method of Doubt*. Descartes believed in the inherent superiority of reason over sensory experiences. *Any* sensory experience could be doubted, but pure deductive reasoning could not, so long as the premises were taken to be true. This mistrust of observational science may occur in unifiers, but does not transfer over to all those we could imagine. The work of James Joule resulted in reducing nature to a few "general principles which will explain everything", i.e. principles for understanding the translation of mechanical energy to heat energy, leaving the universe a little simpler than he found it, yet his approach was critically reliant on rigorous experimental observations.

On the other hand, Bacon was clear on his need for experimental observation, applying the inductive process to produce new knowledge based firmly on the real world. Bacon's view held little space for theories without some experimental grounding – he stated clearly that mathematics should be used *"only to give definiteness to natural philosophy, not to generate or give it birth."* (Novum Organum XCVI, 1620). This places him far from the unifier camp, where abstractions and mathematics are often used to motivate and give birth the next stage of investigation.

A further important aspect of Bacon's philosophy was the *economic impact* of science. Bacon's era was one where the British Empire was emerging, a time of immense economic growth throughout the 16th/17th century. He demands that natural philosophy (science) should be more than merely contemplative, but should be active, put to use to serve the state, not merely hedonistic but should be a platform for business and economic growth. He was a strong supporter that it could provide economic impact in this manner, though he had a long term viewpoint of returns happening in decades rather than short term engineering aims. So, a strong element of Bacon's philosophy appears to *require* economic impact. This seems to be a common diversifier trait, but by no means a requirement; the earlier example of John Wheeler serves to illustrate this, and many of Rutherford's early observations on the structure of the atom had no immediate application.

3.3. The Analytic / Synthetic Distinction

A controversial idea of the past century in philosophy has been the distinction between *analytic* and *synthetic* statements (Kant, 1781). Here we provide a brief discussion, though a fuller treatment is outside the scope of this article. A statement of the form "S is P", is *analytic* if the predicate P is contained within the subject S, that is, the statement *is true in virtue of its own meaning*. The example made famous by Kant was "all bachelors are unmarried" – the term bachelor *means* to be unmarried. On the other hand, a *synthetic* statement is one such as "all bachelors are unhappy", where the predicate is not *necessarily* contained in the subject, and to ascertain its truth requires some information beyond the meaning of the words.

Kant argued for a third category, *synthetic a-priori* statements – here the predicate is not contained within the subject, but the statement is necessarily true and does not require any further information to confirm it as such. Kant asserts that all of mathematics is in this category. If we take an example from our own field – a Kalman filter is a special case of a Gaussian Process. Rephrasing this, we have that a "KF is a GP" – clearly true, and not requiring sensory experience to confirm. Additionally, the definition of a KF in no way uses a GP as a defining component, thus we consider this is a *synthetic a-priori* statement. Whomsoever was the first to notice this (KF=GP) could certainly be regarded as a "unifier" — re-interpreting one Machine Learning model as another, showing how a single principle can unite the two bodies of literature. Could it be that unifiers are more pre-disposed to making synthetic a-priori statements? At present, the answer to this is unclear. Certainly if we take the strictest view of the work we do – it all comes down to a mathematical statement or set of statements (i.e. algorithm) executed on a computer.

Kant (1781) referred to analytic statements as *clarifying or explicating* our knowledge, or in other words, making explicit what was once implicit. Similarly, he referred to synthetic statements as *augmenting* or *extending* knowledge, that is, introducing new information that was not contained in the subject S in any way. These two styles – clarifying versus extending – certainly bear a passing resemblance to the concepts of *unifying* versus *diversifying*. However, to categorise them strictly as such would be to say an act of unifying never extends our knowledge, and Einstein certainly did extend our understanding of the universe. A full in-depth treatment of the analytic/synthetic distinction is outside the scope of this article, but is certainly something worth pursuing in future work.

3.4. *Summary: The Balance between Unifiers and Diversifiers*

Perhaps then, the unifier-diversifier split is best seen as a *spectrum*. The unifiers (e.g. Einstein) are following a direction that bears some similarities to Plato's worldview, though of course modern science bows to experimental tests of its validity – something Plato would never accept. The diversifiers are somewhat Baconian, though without Bacon's strict need for economic impact. The grey area between unifier/diversifier seems to be best characterised by the Aristotlean viewpoint.

If we look at the space of people holding the diversifier mindset, it seems it may be slightly more full of experimenters than theorists, probably slightly more engineers than scientists, and probably slightly more industrial than academic. The space of unifiers is roughly the complement, but none of these individual dichotomies is isomorphic to Dyson's concept.

However, Dyson does perceive a definite imbalance between unifiers and diversifiers in different fields, and this may well hold true for Pattern Recognition also. He argues that the unifiers have dominated physics for most of the 20th century. In contrast, he believes that biology has enjoyed a "healthier balance", where although it is the case that diversifiers have dominated, when a unifier like Darwin or Hamilton comes along, he is not ignored, but celebrated. Dyson reminds us that in biology, Darwin's work is celebrated as a milestone unifying framework, but such occurrences are rare — the working lives of 99 out of 100 biologists consists of investigating and manipulating the complex behaviour patterns of particular species or biochemical pathways. A modern perspective on this divide is given by Jogalekar (2014).

In a controversial point, Dyson states his belief that the unifiers are most likely to be remembered in history:

> *"it is true in general that the very greatest scientists in each discipline are unifiers. This is especially true in physics."*

However, given his later comments, one posits that Dyson intends the term "greatest" here to mean in the sense of fame/notoriety, as opposed to ability or impact. He certainly does not try to downplay the significance of the diversifier stance and, in his own work, engages in both styles of work. He provides an interesting account referring to his work to unify the field of Quantum Electrodynamics:

> *"When I did my most important piece of work* [...] *I had consciously in mind a metaphor to describe what I was doing. The metaphor was bridge-building. Tomonaga and Schwinger had built solid foundations on one side of a river of ignorance. Feynman had built solid foundations on the other side, and my job was to design and build the cantilevers reaching out over the water until they met in the middle."* (Dyson, 1995)

So here, Dyson *explicitly* thinks of himself in a unifying role. In another communication (Dyson, 1979) he recounts his discussions on Quantum ElectroDynamics with Richard P. Feynman, who was apparently obsessed with finding a unifying theory of the large (gravity) and small (nuclear forces). In contrast, Dyson was comfortable with more than one set of equations, each useful at different scales. Referencing Gödel's theorem says:

> *"in the last hundred years of physics, unifiers have had things too much their own way.* [...] *I hope that the notion of a final statement of the laws of physics will prove as illusory as the notion of a formal decision process for all of mathematics. If it should turn out that the whole of physical reality can be described by a finite set of equations, I would be disappointed."*

So here he is quite clear that he also emphasises with a slight diversifier viewpoint, that not everything can be brought under a single metaphorical umbrella. While Dyson clearly thinks that the great advances of 20th century physics are due to the dominant trend of unifiers, he clearly states his final position,

> *"every science needs for its healthy growth a creative balance between unifiers and diversifiers"*

With this more clearly elucidated, we will consider how some of these issues transfer to the Pattern Recognition field.

ACCEPTED MANUSCRIPT

4. Unifying and Diversifying in Pattern Recognition

The previous sections have explored various subtle interpretations of the terms "unifier" and "diversifier". One element made very clear is that unifiers favour *simplicity* in their work. Dyson seems to equate this with a certain *beauty* in the theory or experimental setup – in physics, mathematical beauty is a key element in the pursuit of a Grand Unified Theory for the field. In this section we will address these issues for Pattern Recognition: firstly exploring the concept of beauty, and then the idea of a Grand Unified Theory. Finally, we will consider how these two competing pressures of unifying/diversifying balance against each other over long time periods.

4.1. The Pursuit of Beauty in our Work?

While Dyson does discuss the idea of beauty in theories, several prominent physicists have stated their belief more boldly:

> *"a beautiful or elegant theory is more likely to be right than a theory that is inelegant."* (Gell-Mann, 2007)

In the field of physics, this pursuit of aesthetics has proved exceptionally fruitful. Dyson however does not believe this holds true for all fields of science:

> *"Mathematical beauty was key to the discovery of the laws of nature* [...] *That somehow seemed to work beautifully in physics, but it doesn't seem to work in biology* [...] *The fact is that mathematics is useful for biology only in a very humble way, essentially computer science* [...] *making simulations of complicated systems* [...] *not as a creator for insight."* (Dyson, 2014)

To explain the concept of mathematical beauty is challenging, just as it is challenging to explain the feeling an individual gets from a piece of artwork. The great physicist Murray Gell-Mann said that something is beautiful if it can be explained concisely in terms of mathematics we already have. Richard Feynman explained it as the quality of a result that fits like the last piece in a puzzle, either making everything else seem obvious in hindsight, or providing startling new predictions that are borne out in experiment.

While it seems to be justifiable that in physics, mathematical beauty is the key to truth, in PR we are not necessarily always seeking *truth* — but sometimes simply *utility*. So, is mathematical beauty the key to progress: either to discovery of new truths, or new utility, in our field? Ockham's razor is the obvious discussion point here. The pursuit of simplicity has clearly been a useful practical rule for model selection. In terms of theories/concepts, it has also been a useful post-hoc organisational tool – cleaning up areas after their invention, sometimes yielding small gaps for new work. However the principle has yet to prove its worth at the same magnitude observed in physics, a tool of discovery for entirely new areas of study.

Symmetry is a form of beauty which has been a crucial tool in the understanding of fundamental physics. The most recent high profile example of this is the discovery of the Higgs Boson, predicted to be observable in the Large Hadron Collider at a particular energy level. This prediction was made in 1964, for the simple reason that it would make for a beautiful mathematical symmetry. There may be algorithms we consider beautiful in retrospect – but this principle, of *discovery via aesthetics*, has not yet been so convincingly demonstrated in PR. There has not been a flood of predictions in the form *"there should exist a learning algorithm with generalisation error x%"*. The only instance even vaguely like this (that I know of) is the Boosting family of algorithms — the existence of which was predicted by studies in computational complexity theory (Kearns and Valiant, 1988), and discovered later by Schapire (1990). Our equivalent to the headline-grabbing Higgs prediction would be a prophecy of the form: *"if you create a deep neural network with between* 10^{15} *and* 10^{16} *connections, a phase transition should occur and enable a new level of machine intelligence"*. Thereafter, several billion of EU funding would be directed toward tunnelling under Switzerland to build a neural net big enough. But it has not happened.

It is arguable that a far more fruitful "tool of discovery" in our field has been *inter-disciplinarity*. Many of our best optimisation schemes come from mathematics (e.g. simplex) or physics (e.g. simulated annealing), many of our best models come from biological analogy (e.g. convolutional neural nets), and many of our best methodologies come from statistics (e.g. bootstrap). Of course, this may be an artefact of this stage in our (short) history, and in 50 years the pursuit of symmetry-breaking might turn out useful for artificial intelligence, but who knows.

Given this inherent interdisciplinarity, it is instructive to question whether our subject is on a path toward a 'unified theory' as many people believe is the case for physics. Or indeed, if it is 'unifiable' at all. Even if the answer to this is negative, are there individual elements of our *practice* that could be unified? What are the pros and cons of unifying/diversifying in our field?

4.2. *A Unified Theory of Pattern Recognition?*

Is it the case that there exists a *single* unified theory of Pattern Recognition, toward which we are converging? I believe this idea can be dismissed immediately almost without controversy. For one, even if a Grand Unified Theory exists for physics, we *know* that we are not describing that. It is true that we are in some sense using inference to predict the behaviour of the universe (e.g. whether a person will buy a book on Amazon or not), but we are *modelling* at a level of abstraction several dozen layers above where String Theory is working. And, multiple abstractions can hold true without problem, providing different overlapping and mutually reinforcing viewpoints. The best way to model something is not necessarily at the deepest level at which we understand it. For example, fluid mechanics is a well established discipline, allowing us to predict how water waves break against a wall; the calculations work almost perfectly, using the assumption that the water is a continuum, even though we understand the water to be made of atoms, or digging deeper, little vibrating strings in 11 dimensions. There of course exist almost religious factions that try to convince everybody else that their method of data analysis and inference is the *One True Path*, but ultimately, with incisive questioning, they can usually be brought to a more pragmatic perspective, at least in the short term.

The pragmatic viewpoint says that different mechanisms, theories and implementations of intelligence will be useful in different scenarios. In this light, it can be believed that we may get pockets of unification, but no over-arching theory to unify us all. There is certainly no shortage of attempts — a quick Google search reveals recent papers in a common style:

- A Unifying Framework for Statistical Relational Learning
- Rule Evaluation Measures: A Unifying View
- A Unifying View of Multiple Kernel Learning
- A Unifying Framework for Information Theoretic Feature Selection

Just as Dyson believes that different sets of equations would be useful at different scales of experience, so it is likely that different theories of inference and data modelling will be appropriate for different problem scenarios, and at different scales and types of data, different aesthetics will become apparent.

4.3. *Unifying/Diversifying as Part of an Evolutionary Process*

In any field, unifying and diversifying behave according a kind of evolutionary process[3], where the units of evolution are *memes*: ideas, behaviours or styles that spread from person to person within a culture. In PR, the Bayesian and Kernel memes gained particular traction from about 2000 onwards, and are arguably of a *unifying* flavour. They have both successfully abstracted several techniques to special cases of their respective methodologies, enabling new insights, e.g. kernel PCA. Whilst these have proved immensely powerful, when a unifying meme does not serve to progress science as rapidly as it has in the past, its dominance in the culture is displaced, and diversifying memes appear. The recent meme of *deep learning* seems to be very much in a diversifying flavour, without a single aesthetically pleasing theory to explain it, yet clearly providing results of utility in several domains.

The evolutionary pressures are complex, existing at various *levels of granularity* in a field. Coarse-grained aspects of the field can be *unified* under a common societal challenge, such as the recent trend for *Big Data*, while at the same time finer-grained aspects of the field are *diversified* to cope with the new challenges. Sometimes results are rediscovered, but put in a new light given that other topics have progressed in the meantime — results once seen as diversifying, can later be seen to serve a unification. The field progresses only with the *blend* of these pressures: too much of either one will hinder progress.

Taking this long-term viewpoint, a meaningful diversification can be exceptionally healthy. The field of artificial intelligence underwent the ultimate diversification from the 1980s onward, documented well by Cristianini (2014). Looking back at the Dartmouth Conference

[3]Thanks to Fabio Roli for suggesting this metaphor.

(McCarthy et al., 1955), the primary goal was to create intelligent beings, and it was imagined this was only a decade or so away. Over half a century later, we have sub-sub-subfields—for example, *adversarial* classification, as a subfield of supervised learning, as a subfield of machine learning, as a subfield of AI. This diversification enabled questions we never imagined 1955, and created a generation of technology that has become indispensable to everyday life.

5. What do unifiers/diversifiers do in PR?

Unifiers like to explore the connectedness of ideas. They prefer to discover the relationships between existing scientific artefacts, rather than create new ones. In our case, these are algorithms, mathematical constructions, and their implementations, whether in hardware or software. When unifiers create, they create artefacts at the intersection of existing ones, so as to see their connection. As a consequence, they write papers which bring people together, cross discipline boundaries for the purpose of reinterpreting their primary field of study, and have a broad view of the research landscape. It could be conjectured that unifiers tend to publish less frequently than diversifiers, taking more time to integrate the various concepts they bring together.

Diversifiers on the other hand enjoy exploration and invention, they have a narrower focus on the research landscape at any one time, or multiple narrow foci. They push limits and figure out what is and is not possible. This was precisely Rutherford's achievement, pushing the boundaries of our understanding of the atom. As a consequence, they innovate more, posing questions not previously considered, often by importing ideas across a discipline boundary. They create artefacts with utility, not only in the immediate engineering sense but also in that they highlight problems, chinks in the armour of a theory.

Both groups have the capability to inspire new directions, create new fields— but they do it in different ways. Unifiers provide a new viewpoint on existing literature, showing gaps, enabling meaningful analysis of properties, providing new languages which can express computational artefacts at the junction of several others. Diversifiers address challenging new domains and questions that current work has simply not considered; they start slow, often with heuristic (but effective) approaches and accumulate a fan-base of loyal followers who slowly refine these, figuring out what works and what does not. As may be obvious, the two groups provide fuel for each other—one follows the other in a never-ending cycle of interleaving innovation. Where the unifier defines a framework, the diversifier finds an exception. When the diversifier finds sufficiently many exceptions, the unifier sees commonalities and patterns for a new framework.

As we have stated at the start of Section 2.3, unifying and diversifying are *styles of practice*. As such, each can be done *badly*, causing more harm than good. Whilst it is tempting in this essay to take the middle ground and avoid offence, here I will not, and outline downsides of each practice.

Novice unifiers[4] can be dangerous. They often stumble for a long time, seeing patterns where there are none, over-egging the significance of their 'frameworks'. Unification can be *sterile* — bringing several ideas under a common umbrella, but ending up with strained analogies and relationships between the ideas, and ultimately closing more doors than it opens. Claims to unification can be little more than a *categorisation* of the ideas: a literature review with a solid backbone, but not enabling invention of new ideas, or meaningful explanation of existing ones. There are downsides to the unifier stance in general, even if done "correctly". Why should we force others to adopt the same perspective as our own? If we attempt to cast everything into a single mould then, whilst aesthetically pleasing, it will mean compromises have to be made. In this way, doors will be closed on young minds exploring the literature for the first time.

As stated eloquently by Langley (1989), too much diversification can also be bad for a field:

> *"...diversification also has its dangers. Subdisciplines can emerge that focus on one goal or evaluation scheme to the exclusion of others, and similarities among methods can be obscured by different notations and terminology."*
> (Langley, 1989)

It is common to see papers offering 'novel' methods with an *immensely* complex computational pipeline, and many

[4]The author confesses to being in this category for a good while.

parameter settings left unjustified; or worse, nuances of the implementation left completely unreported. Though this article is focused on Pattern Recognition, it is fair to note that similarly vague work appears in related communities (Sörensen, 2013). These papers can generally be characterised by the phrase *"my classifier gets higher accuracy than your classifier"*, though results often cannot be reproduced as they depend on those unspecified nuances. The best one can say in this situation is congratulations to the authors for finding the three or four datasets on which their method was successful. It is difficult to see what can be done to discourage this, apart from refining the unspoken rules of acceptable practice in our field. One can only hope that further standards of *reproducible research* will infiltrate the community, and allow genuine progress rather than illusory (Hand, 2006).

6. Who Drives the Paradigm Shifts of Pattern Recognition?

Kuhn (1970) presents a treatise on the nature and reasons behind revolutions in scientific understanding. In this, he discusses how new scientific concepts bring about revolutions in a field. Kuhn proposes that all scientific revolutions follow a similar pattern, described by his 'paradigm shift' cycle.

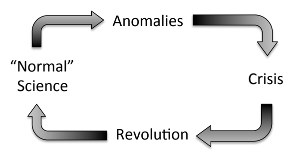


Figure 1: Kuhn's paradigm shift cycle. So-called 'normal' science precedes anomalies in observation, followed by a crisis of understanding, then a scientific 'revolution' where new ideas are adopted by mainstream science, and a new paradigm begins.

An interesting question is *who drives the transitions around this cycle?*. Is it the unifiers, coming up with fundamental new theorems to unify the state of the art? Or the diversifiers, challenging popular opinion with new observations / phenomena?

Dyson claims that Kuhn's vision of this is too narrow - that transitions are brought into being only by unifiers, coming up with new theorems. He calls this a *concept-driven revolution*. He expands upon this view in a book titled *"The Sun, The Genome, and The Internet"*, (Dyson, 1999), discussing how *tools* are an equal (if not greater) influence on the recent revolutions in science. Dyson takes a very broad viewpoint on the definition of a tool, as might be suggested by the book title, he considers the sun, the genome, and the internet, all as tools for science. In his own words:

> *"a scientific tool is not only considered to be something that strengthens our senses or is useful in taking measurements, but also as an aid to our understanding"*
> (Dyson, 1999, p51)

Dyson points out that new tools (created by diversifiers) enable observation of new phenomena, which possibly conflict with previous theories — pushing the field into the 'anomalies' and then 'crisis' phase. In addition, the transition back round to normal science is very often enabled by tools which observe and manipulate data to resolve the anomalies — the new theories play a relatively minor role in the process. He proposes that physics was dominated by concept-driven revolutions prior to the 20th century, but beyond the 1920s it was not possible to conduct experiments in isolation, and tool-driven revolutions took over — when tools like Electron Microscopes and the Large Hadron Collider enabled new paradigms of understanding.

Returning the discussion to Pattern Recognition, we had a concept-driven revolution in the early 1990s, when statistical and data-driven modelling began to dominate. We are possibly about to transition into a *tool-driven* revolution, with the availability of tools like Kinect, and a new wearable computing industry with the Apple Watch and imitators – using many embedded sensors that will make inferences based on observations during the day. Embedded sensors and computing are disappearing into the fabric of life, as so many Sci-Fi movies have it. This generation of embedded intelligent sensors will need software, and it will be the intellectual descendants of those reading this article that write it. Other non-obvious examples of ML tools are: approximate inference algorithms, multi-core computers and GPUs, Amazon's Mechanical Turk, toolkits such as Weka and libSVM, and of course APIs that open up sources of data unavailable to most people, such as Twitter feeds.

Very occasionally, single, powerful tools come along that drive an entire revolution. These enable the field to both ask and answer questions that would have been previously inconceivable, just as the physics of black holes would be inconceivable to the ancient Athenians without modern day radio telescopes. Dyson has recently speculated (Dyson, 2014) that the progress of artificial intelligence is fundamentally limited until critical new tools (analogue computing machines) are properly developed.

7. Related Work in Pattern Recognition

Whilst we may never have a truly "unified theory of inference", there are a number of technical elements of our field which could benefit from a little unification; in classic papers, Breiman (2001) and Langley (1989) present ideas along this line. With over a decade since Breiman's paper, and a quarter-century since Langley's, it seems an interesting time to revisit their words.

7.1. Breiman's Two Cultures

Breiman (2001) discusses two cultures of statistical modelling: *data* modelling versus *algorithmic* modelling. For Breiman, "data modelling" means considering the form of the problem/data one is faced with, then thinking of a parametric class of mathematical models, and fitting the parameters. This is epitomised by linear and logistic regressions, and procedures like LASSO. The models are mathematically tractable and elegant, and have direct (if questionable) mappings of their structure to phenomena in the problem domain – Breiman proposes that their use accounts for 98% of the working lives of all statisticians. The 'algorithmic modelling' culture is epitomised by decision trees, neural nets, SVMs and other terms familiar to us in Machine Learning. These models make no claim to reflect the structure of the problem in their own structure. A node in a neural net or an RBF kernel is simply a good way of fitting the data, as opposed to being a symbol for a particular real-world event. He claims that these account for 2% of all statisticians.

Breiman argues three main points, that the overuse of data models has: led to irrelevant theory and questionable conclusions; kept statisticians from using more suitable procedures; and kept statisticians from working on exciting new domains.

In the years since Breiman's paper, many of these boundaries have been crossed. Many of the techniques Breiman calls algorithmic models are now known as computational statistics, and are in common use in both communities. There is still the hardcore of both the statistics and ML communities that hold fast to the aesthetics of certain modelling approaches, and these are slowly dropping their restrictive assumptions, becoming just as strong in practice as any method. Nevertheless, Breiman's paper is a thought-provoking read over a decade later; and, with its pragmatic view, encourages these two communities to unify their goals and practices.

7.2. Langley's Seven Dichotomies

Langley (1989) wrote a striking editorial for an issue of Machine Learning Journal, on the topic of unifying machine learning as it stood in the late 1980s. He discusses seven apparent dichotomies that had emerged at the time, stating that *"long term progress will only occur if we can find ways to unify these apparently competing views into a single whole"*.

For the interested reader, an online Appendix[5] to this article contains a detailed analysis of each of these seven dichotomies (from my own perspective) and the extent to which I believe they have been resolved over the past 25 years.

8. The Value of the Dichotomy

What is the *value* of Dyson's unifier/diversifier dichotomy? Is it *predictive*, in the sense that it may inform profitable new directions in research? Does it give an understanding of our own actions, successes/failures? What *is* a healthy balance of the two for PR? To address these questions, it is in fact easier to first address the more abstract question — what is the value of philosophy in general? A powerful answer to this is presented by Russell (1912, chapter XV).

> *"The study of physical science is to be recommended* [...] *because of the effect on mankind in general. This utility does not belong to philosophy. If the study of*

[5] `http://www.cs.man.ac.uk/~gbrown/research/langley.pdf`

philosophy has any value [...] *it must be only indirectly, through its effects upon the lives of those who study it."*

Russell suggests that the benefits of philosophical thought show themselves firstly in the lives of those who study it, and only indirectly in the lives of others who they interact with. If philosophical reflection is undertaken by scientists, this enables more considered and reflective practice in whatever field they happen to occupy. Many years later, Russell summarised his views as,

> *"the man who has no tincture of philosophy goes through life imprisoned in the prejudices derived from common sense, from the habitual beliefs of his age or his nation, and from convictions which have grown up in his mind without the co-operation or consent of his deliberate reason."*
> (Russell, 1946)

Can philosophy in general inform profitable new directions in PR research? Pelillo and Scantamburlo (2013) present a post-hoc example, discussing the *essentialist* viewpoint in philosophy. Essentialism is the view that entities in the world have inherent, *essential* and *immutable* properties, by which they can be described. Pelillo and Scantamburlo (2013) discuss how *dissimilarity measures* in PR sit in direct opposition to this philosophical view, in that an entity is best described by its relation (similarities) to other entities. I suggest an speculative extension, that *evaluation metrics* in our field should be considered in the same manner. The evaluation metric is not something *essential* to a predictive model — it is something subjective, imposed by humans, for particular tasks, at a particular point in time, in a particular context. Perhaps merging some ideas from multi-objective optimisation with the theory of ranking measures might prove fruitful in this direction.

But, what benefit is the unifier/diversifier dichotomy? Can it give an understanding of our own actions, successes/failures? Possibly, yes. This may come by reflections on previous successes/failures, and what character they have. If an individual sees on reflection that most of their successes have come from a diversifying mindset, they may be able to direct actions accordingly in the future. As in most areas of life, raising awareness of one's emotions and actions is usually profitable.

What is a healthy balance of unifying and diversifying? This is a difficult question. For one, a 'healthy balance' is ill-defined, and even if it wasn't, it would be different at different scales of organisation. The individual researcher, their research team, their academic school or institution, all the way up to a nation's research budget — each will have different emphases depending on complex factors. Even so, the dichotomy might be useful as an analytical tool — asking at each of these scales, how much effort is being placed into four different areas: unifying in science/engineering, and diversifying in science/engineering. Whilst this is sure to be a controversial point of view, and I do not suppose to know the best way to manage an entire country's research budget, it is a thought-provoking concept.

9. Conclusion

The question originally posed for this paper was "can Dyson's unifier/diversifier dichotomy apply in the Pattern Recognition field?". We conclude that the answer is *yes*, but in a subtly different way. There do exist *pure* unifiers and *pure* diversifiers, just as in physics, however, they are rare in our field. Most people in PR sit on the spectrum between the two: keen to expand the scope of our field (diversifying) but equally keen to find aesthetically pleasing results linking them to other researchers (unifying). Many researchers can adopt unifying perspectives one day, and be diversifying the next.

However, I conjecture that this situation may change with time. Dyson argued that both theories and tools drive scientific revolutions (Dyson, 1999). Physics has been maturing its arsenal of theories and tools for hundreds of years longer than us. As such, the *accessibility* of theories and tools is much more restrictive than in PR. It takes many years to master advanced physics, such as the mathematics of String Theory. Equivalently, to be an experimental physicist, tools such as electron microscopes cost hundreds of thousands of pounds; or, even more extreme, the Large Hadron Collider cost billions of Euros, and is inaccessible to most. In biology, an experimentalist spends many years learning just one tricky technique for refining biochemical reactions to observe their phenomenon of interest. In Pattern Recognition however, we all have Matlab, Python, Weka, and fast computers, the tools we need to do good work. The theory side is also not too obscure — a student in computational learning theory can start to contribute excellent work before they graduate. It is likely this will change with time as the field matures.

Experimental tools are maturing too. GPUs and multi-core machines are commonplace, but learning how to program them is non-trivial. Datasets such as Twitter-feeds are posing new challenges to us, and dealing with this data scale takes special skills, and increasingly, large budgets. Datasets too may become commodities; we are already seeing that only large industry practitioners can afford tackle certain types of problems. Take it further into the future – what happens when the tools of Machine Learning are custom built neural microchips, ala the *positronic net* of Commander Data in Star Trek? This is not such an implausible direction to head. Why should we not have computing devices that are by construction inherently suited to inference tasks, either deductive or inductive? It could be argued that such devices are at prototype stage already with *neuromorphic* computing (Furber et al., 2014).

As the theories in the field becomes more sophisticated, the techniques will take many years to learn. As the tools become more specialised, they will become more financially inaccessible. Although we have argued that theory/experiment is not an isomorphic dichotomy to unifier/diversifier, it is one factor. If theorists and experimenters become very distinct roles in PR, with theorists rarely (if ever) learning the tools of experimenters, and vice versa, then the gulf between unifier and diversifier may grow larger.

So, maybe we will never have a fully unified discipline of Pattern Recognition. It is likely however that we need *both* the forces of unification and diversification to move forward, summarised eloquently by Langley (1989), referencing Dyson:

> *"Just as the twin forces of gravity and pressure hold a star in dynamic equilibrium while generating energy, so the joint processes of diversification and unification can hold a science together while fostering progress."*

To conclude, in my own career to date, I've mostly been a unifier. However, Dyson defines diversifiers as those who like exploring the details of nature, and that unifiers prefer the broad brush, the big picture. I believe fruitful unifications only come from looking at the details — so maybe I'm a bit of both, but it's fun to consider.

Acknowledgments

I would like to thank Fabio Roli for prompting and encouraging the thought processes that led to this work, and Joshua Knowles, Marcello Pelillo, Michael Lee and Henry Reeve for broadening my awareness of the philosophy literature. The research leading to these results has received funding from the EPSRC Anyscale project EP/L000725/1.

References

Breiman, L., 2001. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). Statistical Science 16, 199–231.

Cristianini, N., 2014. On the current paradigm in artificial intelligence. AI Communications 27, 37–43.

Darwin, C., 1887. The Autobiography of Charles Darwin. J. Murray, London.

Dyson, F.J., 1979. Disturbing the Universe. Basic Books.

Dyson, F.J., 1988. Manchester and Athens, in: Infinite in All Directions. Harper & Row Perennial, pp. 35–53.

Dyson, F.J., 1989. Roundtable Critique; Hidden in Plain: View, Narrative and Creative Potentials in War and Peace. Journal of Tolstoy Studies 2, 1–41.

Dyson, F.J., 1995. The Scientist as a Rebel. New York Review Books.

Dyson, F.J., 1999. The Sun, The Genome, and the Internet: Tools of Scientific Revolutions. Oxford University Press.

Dyson, F.J., 2014. Are Brains Analogue or Digital?, School of Theoretical Physics, Dublin Institute for Advanced Studies, 19th May 2014.

Feynman, R.P., 1964. The Character of Physical Law (Messenger Lectures).

Furber, S., Galluppi, F., Temple, S., Plana, L., 2014. The SpiNNaker Project. Proceedings of the IEEE .

Gell-Mann, M., 2007. TED Talk: Beauty, Truth and... Physics?

Hand, D.J., 2006. Classifier technology and the illusion of progress. Statistical Science 21, 1–14.

Jogalekar, A., 2014. Diversifiers of the world - Unite!, in: Scientific American Blogs, April 15th 2014.

Kant, I., 1781. Critique of Pure Reason, translated by N. Kemp-Smith (1929).

Kearns, M., Valiant, L.G., 1988. Learning Boolean formulae or nite automata is as hard as factoring. Technical Report TR-14-88. Harvard University Aiken Computation Laboratory.

Kuhn, T., 1970. The Structure of Scientific Revolutions. Univ. Chicago Press.

Langley, P., 1989. Toward a unified science of machine learning. Machine Learning 3, 253–259.

McCarthy, J., Minsky, M., Rochester, N., Shannon, C., 1955. The Dartmouth Summer Research Conference on Artificial Intelligence. URL: `http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html`.

Pelillo, M., Scantamburlo, T., 2013. How mature is the field of machine learning?, in: Advances in Artificial Intelligence. Springer, pp. 121–132.

Pelillo, M., Scantamburlo, T., Schiaffonati, V., 2015. Pattern Recognition Between Science and Engineering: A Red Herring? Pattern Recognition Letters (this special issue) .

Russell, B., 1912. The Problems of Philosophy. Oxford University Press.

Russell, B., 1946. Philosophy for Laymen. Universities Quarterly 1, 38–49.

Schapire, R.E., 1990. The Strength of Weak Learnability. Machine Learning 5, 197–227.

Sörensen, K., 2013. Metaheuristics: the metaphor exposed. International Transactions in Operational Research, DOI: 10.1111/itor.12001 .

Thackray, A., 1974. Natural Knowledge in Cultural Context: The Manchester Model. American History Review 79, 672–709.