

Revisoning the unification of syntax, semantics and statistics in shape analysis

Anand Rangarajan^{1,*}

Department of Computer and Information Science and Engineering (CISE), University of Florida, Gainesville, FL, 32611-6120, USA

Abstract

Hidden patterns, rendered visible by hindsight, often stand revealed as strong influences. Rama Chellappa's lab in the mid to late '80s molded our character and scholarship in more ways than one. Rama ably handled the transition from image processing to computer vision and established an applied math and computing infrastructure from which we continue to benefit. In particular, the themes important to King-Sun Fu—*syntax, semantics and statistics*—were all debated in Rama's lab at that time. We argue that this triad remains important. With syntactic representations losing mindshare to statistics, we remain in the hunt for unification. And with the syntax versus semantics debate unresolved, it deserves a hearing as well. We offer two themes—uncertainty and interaction—to aid in the process of unification. First, we show that complex wave functions carry probabilistic location information in their magnitude and syntactic (curve) information in their phase while representing uncertainty at a fundamental level. Next, after reviewing work in analytic philosophy, we connect semantics to intentional, mental content. Analytic philosophy reminds us to take human experience seriously but remaining physicalist if possible. To this end, we introduce a nondualist interactionist model of experience, wherein compositional (physical) subjects are constantly shaping and being shaped by a physical world. We then demonstrate that wave functions can accommodate

*Corresponding Author

Email address: anand@cise.ufl.edu (Anand Rangarajan)

¹This work is partially supported by NSF IIS RI:1143963.

interaction, closely tracking previous work in physics on the measurement problem. The linearity and superposition properties of wave functions allow for literal addition of waves created by human interaction with shapes. Finally, we briefly survey the current situation in the human-computer interaction (HCI) field and argue that mathematical models of interaction akin to those in pattern recognition can aid HCI. We close by arguing that we can follow in Fu's footsteps and incorporate the mathematical modeling of human interaction into pattern recognition.

Keywords: syntax, semantics, statistics, uncertainty, interaction, wave functions, Schrödinger, Hamilton-Jacobi, phase, topology, level-sets, distance transform, intentionality, qualia, compositionality, HCI

1. Invocation

Hold the flame 'til the dream ignites

A spirit with a vision is a dream with a mission

Rush, Mission

Luck is often described as “being in the right place at the right time”. Looking back, it seems clear that we were very fortunate indeed to inhabit Prof. Rama Chellappa's lab during the mid '80s, in the Signal and Image Processing Institute (SIPI) at the University of Southern California (USC) set amidst the backdrop of a sophisticated, urban Los Angeles. Computer vision was in its infancy, struggling to emerge from the shadow of an established artificial intelligence (AI). Pattern recognition was seeking to reinvent itself in the hands of a resurgent neural networks field [1] and would subsequently find a more stable partner in machine learning [2]. If there was interest in human centered computing, we certainly didn't see much evidence at that point.

Rama was and remains the embodiment of that transition from a field taxonomy so stable (in the late '70s) that it verged on taxidermy to a period (the

early '80s) when “all that is solid melts into air”² [3]. Not content with resting on his laurels in signal and image processing, Rama sought to bring the clarity and rigor found in mature (and older) fields that stood on the shoulders of applied mathematics to the nascent (and therefore fertile) area of computer vision. This quest was aided by three significant factors: (i) the brilliance of Marr [4] in clearly articulating the nature of representation in computer vision, (ii) the integrative genius of Fu [5] in bringing together syntax, semantics and statistics in pattern recognition, and (iii) the strong reliance of neural networks on statistical mechanics, nonlinear optimization and applied mathematics in general. While we return to these themes frequently in this work, it is to the specifics of Rama’s contributions—and the manner in which they illustrate the deployment of applied math in computer vision—that we now turn.

Since we’re using a signal processing to computer vision crossover perspective, we bypass Rama’s considerable work on Markov random fields in image processing and analysis [6] except to note in passing that this framework has served us well over the past three decades. Instead, we highlight the twin contributions in shape from shading and motion estimation to buttress our crossover points. In shape from shading, the problem of enforcing integrability was a thorn in the variational framework. Frankot and Chellappa [7] designed an elegant projection algorithm to find the closest, valid surface in a least-squares sense to the non-integrable surface obtained via the calculus of variations approach. The projection method was signal processing inspired through and through using the fast Fourier transform (FFT) and least-squares estimation to full effect. Next up, in motion analysis, Broida and Chellappa [8] were the first to use a Kalman filter to recursively estimate rigid body motion parameters from noisy images. Once again, the application of signal processing methodologies (filtering, estimation etc.) bore immediate fruit in an important computer vision problem. In Rama’s lab, these early successes reinforced in all of us not only the impor-

²The first IEEE Computer Vision and Pattern Recognition (CVPR) conference was held in 1983.

tance of mathematical rigor in vision problems but the impact of ideas drawn from other fields such as signal processing on the fledgling area of computer vision. Furthermore, the specific successes in shape from shading and motion estimation emboldened Rama to emphasize (i) the correspondence problem in object recognition and motion and (ii) the link between variational problems and specific differential equations such as the Poisson equation [9, 10]. The emphasis on correspondence in the face of an ascendant optical flow paradigm was particularly courageous and key to shaping my views [11, 12] even though I didn't realize it at that time.

SIPI culture in general and Rama's background in MRFs in particular meant that inference and estimation principles were drilled into us usually via statistics courses taken in the math department. Geman and Geman's landmark paper on MRFs [13], Gibbs sampling and simulated annealing forced us to pick up quite a bit of statistical mechanics as well. Entropy, computational temperature and more broadly, the role of uncertainty became common themes in conversations. The change in worldview accompanying the transition from image processing to computer vision meant that we had to grapple with artificial intelligence for the first time. For many of us with undergraduate engineering backgrounds—with a lack of emphasis on philosophy—this was simultaneously uncomfortable and exciting. In particular, this implied taking the central issue of representation in AI seriously. Syntax versus semantics [14], the Chinese room [15] and more generally the nooks and crannies of strong AI [16] were endlessly debated. It is no coincidence that Fu's seminal work on the unification of syntax and semantics [5] with a strong focus on statistical pattern recognition became a linchpin. Fu forced us to focus on the triad—*syntax, semantics and statistics*—while remaining grounded in the nuts and bolts of pattern recognition and computer vision.

We argue that Fu's triad is even more relevant today. While the past thirty years has seen much progress in the development of robust algorithms (for classification and regression), it is fair to say that the dominant paradigm has been statistics with syntactic representations frequently discarded or sidelined due

to perceived (and actual) brittleness in the latter. The importance of uncertainty in representation has also faded and replaced (unsuccessfully as we shall argue) by variance and entropy in parametric and non-parametric probability distributions. The syntax versus semantics debate has not fared any better in mainstream AI circles. While analytic philosophy has made tremendous strides in the past twenty years in delineating the distinction between the computational and the experiential, these advances have not been absorbed in information processing circles. Concomitantly, we have seen the rise of human centered computation and the field of human-computer interaction (HCI), but we are not aware of any serious attempts to educate the (usually) younger denizens of HCI with the significance of Fu’s triad in their endeavors. Consequently, HCI and related areas remain unaware of the importance of interaction (between the experiential and the computational) in anchoring semantics while informing syntactic and statistical representations. To this end, we identify the twin themes of uncertainty and interaction as central and attempt to unpack their significance in this essay.

2. Uncertainty

*Who can face the knowledge
That the truth is not the truth
Obsolete
Absolute*

Rush, Distant Early Warning

In §1, we introduced Fu’s contributions in the larger context of the entire field of pattern recognition. In this section, we focus on shape analysis in our attempt to discuss uncertainty in syntactic and statistical representations. We will not be concerned with semantics here, which is deferred to §3.

Representations inspired by probability theory and statistics—placed under the same rubric here—have thrived and prospered in shape analysis. Especially when shapes are parametrized by point-sets, probabilistic representations have

become quite popular due to the relative ease of density estimation in lower dimensions. In the past twenty years, shape correspondence, non-rigid deformable matching, shape dictionaries etc. have all seen considerable progress since the robustness afforded by the representations has allowed for outlier detection, incomplete shape matching and so on. In many cases, shape density functions are first estimated using Parzen windows or related methods. Subsequently, shape densities are matched using entropy minimization or other criteria to obtain shape deformation, shape atlases [17] and the like. Note the simplicity of the representation schemes which feature little to no explicit syntax. Point-sets are generally i.i.d. allowing for straightforward density estimation and the Hausdorff topology prevents the use of relational information. In spite of this, the robustness of the representation can account for both point jitter and outliers belying the need for synthesizing relational and statistical information.

When shapes are parsed into sets of non self- or other-intersecting closed planar curves (in 2D), level sets and distance transform representations have become popular in this space [18]. Distance transforms satisfy the eikonal equation $\|\nabla S\| = 1$ (with a constant forcing function) with the zero level sets comprising the shape. In contrast to the point-set representation, distance functions embed curve syntactic information into a scalar field $S(x)$. There is no room for uncertainty in the representation however, for signed and unsigned distance functions are highly constrained geometric objects that leverage the curve (relational) topology information. These constraints (implicit in $\|\nabla S\| = 1$) do not allow distance functions to be added, for example, or facilitate shape atlas computation. When one shape contains two curves and another three curves, their scalar distance transform fields cannot be easily combined. Consequently, while this representation has flourished with active contours and level sets seeing wide usage in shape extraction, there has not been much interest in these (more syntactic) representations for computing shape statistics.

In the past few years, we have been exploring the notion of uncertainty in the distance transform representation [19, 20]. The eikonal equation $\|\nabla S\| = 1$ is an example of a static Hamilton-Jacobi equation. In the theoretical physics

literature, it is well known that the nonlinear Hamilton-Jacobi equation can be embedded in a linear Schrödinger equation [21]. Since we cannot expect much familiarity with this literature, we summarize the main results below. Assume a particle with mass m moving under a constant potential $V(x)$. The Schrödinger equation for the free particle is

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(x)\psi \quad (1)$$

where \hbar is the reduced Planck's constant. Since $\psi(x)$ is complex, we may express it as $R(x) \exp\left\{i\frac{S(x)}{\hbar}\right\}$ with $R(x)$ being the magnitude and $S(x)$ related to the phase of the wave function. The complex Schrödinger equation can be rewritten as two equations—one for $R(x)$ and the other for $S(x)$:

$$2R \frac{\partial R}{\partial t} + \frac{1}{m} \nabla \cdot (R^2 \nabla S) = 0, \quad (2)$$

$$\frac{\partial S}{\partial t} + \frac{\|\nabla S\|^2}{2m} + V(x) - \frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} = 0. \quad (3)$$

The equation for $S(x)$ is very close to the classical Hamilton-Jacobi equation for a particle. As $\hbar \rightarrow 0$, and assuming bounded derivatives and $R(x) > 0, \forall x \in \Omega$, we get

$$\frac{\partial S}{\partial t} + \frac{\|\nabla S\|^2}{2m} + V(x) = 0, \quad (4)$$

the Hamilton-Jacobi equation for a particle moving under $V(x)$. When we specialize to the static case ($\frac{\partial S}{\partial t} = -1$) and set $m = \frac{1}{2}$ and $V(x) = 0$, we get $\|\nabla S\| = 1$, the constant forcing function version of the eikonal equation. We can solve the linear Schrödinger instead of the nonlinear Hamilton-Jacobi—a concrete payoff of this relationship. Furthermore, the use of a computational parameter similar to Planck's constant introduces uncertainty into the distance transform setup. Finally, instead of computing distance functions, we compute wave functions which by virtue of satisfying a linear differential equation, inherit the properties of linearity and superposition. For more details, please see [22, 20].

What we gain—the introduction of uncertainty and the ability to add and superpose wave functions—more than makes up for what we lose—certainty in

the distance function representation. Having discussed the benefits of a wave function representation over those of a distance function, we now demonstrate one of its advantages—specifically the ability to represent a set of planar curves with uncertainty—relative to shape density functions which were previously used to represent point-sets with uncertainty.

Consider a set of points $x_i \in \mathbb{R}^2$, $i \in \{1, \dots, N\}$. We may fit a density function to this point-set by employing a Parzen window density estimator and a Gaussian kernel. If a Gaussian is centered at each point, we get

$$p(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{\|x - x_i\|^2}{2\sigma^2}\right\}, \quad (5)$$

where σ is the isotropic standard deviation of each 2D Gaussian. The uncertainty parameter σ in this point-set density function representation can be estimated via cross-validation or other methods available in the density estimation literature. The scalar field $p(x)$ gives us the probability density at any location in \mathbb{R}^2 . A large value of the density translates to higher likelihood of the shape being present and vice-versa. As mentioned earlier, there is however no way to represent curves (other than as discretized point-sets) in this framework.

Can we translate the density function uncertainty representation of a point-set to a wave function uncertainty representation of a set of curves? The analysis above suggests that the magnitude of the wave function $R(x)$ can be related to the shape density whereas the phase of the wave function is related to the shortest distance $S(x)$ to the shape boundary. To drive this metaphor home, we turn to Huygens’ principle [23]: we consider waves emanating from every point on the shape boundary. The influence of each wave isotropically decreases as you move away from its source location and the directionality of the wave depends on the normal vector at the source location. From a computational perspective, we cannot place a wave source at each point on the shape boundary—instead we consider planar curves with a discrete number of source locations.

Assume a set of planar non self- or other-intersecting curves parametrized by $x_{kl}, \nu_{kl} \in \mathbb{R}^2$, $k \in \{1, \dots, N_l\}$, $l \in \{1, \dots, L\}$. Here x_{kl} , $k \in \{1, \dots, N_l\}$ is the set of control points of the l th curve in anti-clockwise order and ν_{kl} , $k \in$

$\{1, \dots, N_l\}$, the set of outward facing unit normal vectors. We use unit normals since Huygens' principle only needs directions to be specified. We unpack Huygens' principle by constructing the wave function

$$\psi(x) = \sum_{l=1}^L \sum_{k=1}^{N_l} \exp \left\{ -\frac{\|x - x_{kl}\|}{\lambda} \right\} \exp \left\{ i \frac{\langle \nu_{kl}, x - x_{kl} \rangle}{\tau} \right\}, \quad (6)$$

where λ, τ are free parameters related to isotropic decay and spatial frequency respectively and $\langle \cdot, \cdot \rangle$ denoting an inner product. (The choice of absolute rather than squared distance in the exponent is not justified here.) The wave function does not need to be normalized. The magnitude and phase of the wave function can be computed from (6) and we focus on the phase following its approximate identification with the Hamilton-Jacobi field $S(x)$ above:

$$S(x) = \arctan \left(\frac{\Im \{\psi(x)\}}{\Re \{\psi(x)\}} \right), \quad (7)$$

is the wrapped phase component of the wave function with its zero level set periodicity proportional to τ . Computation of the wrapped phase is straightforward (while certainly not the case for the unwrapped phase). Below, in Figure 2, we display the contours of $S(x)$ corresponding to a set of shapes depicted in Figure 1. For each of the silhouettes³ in Figure 1, we extracted the contours and then fit outward facing normals using B-splines. Then, we compute $\psi(x)$ using the normals along with the curve locations. For all six shapes, we set λ to 0.01 and τ to 0.4 (except for the horse shape where τ was set to 0.3). The level sets of S were extracted via the `contour` function in MATLAB[®]. In *all* cases, we observe (and the reader can observe this as well by zooming into the plots) that the phase information captures the entire *figural* of the shape. Shape topology information is preserved and in particular, we clearly see the medial axes. The reader may be doubtful of the usefulness of the topology information since we displayed all contours of S but this can be dispelled by visualizing the magnitude of the wave function. This information, available in Figure 3 clearly demonstrates the preservation of shape location information in the form of high

³We thank Kaleem Siddiqi for providing us with the set of silhouettes.

magnitude values near the shape boundary.

To the best of our knowledge, this is a new contribution. While the magnitude of the wave function in (6) can be expected to behave similarly to the point-set Parzen window density estimator in (5), topology representation via the phase of the wave function—obtained by merely turning curve normals into spatial frequency attributes—is new. Syntax and statistics are joined at the hip in this new representation: witness the use of attribute information (curve normals) to anchor a relational (topology) representation. Uncertainty in this representation is at a deeper level than density functions since these can be obtained from wave functions and not vice-versa. It remains to be seen if the wave function shape representation can be leveraged for shape atlases, registration and indexing—and we plan to pursue these topics in the near future.

3. Interaction

Nothing he's got he really needs

Twenty first century schizoid man

King Crimson, 21st Century Schizoid Man

Fu's main concern was the unification of structural (syntactical) pattern recognition and statistical decision theory. In [5], he proposed an attributed grammar for pattern recognition. He argued that pattern primitives could be characterized using feature vectors (and their attributes) and handled using statistical decision theory while contextual information could be handled using syntactic (relational) methods. Fu uses semantics in different senses of the word. He sometimes appeals to attribute information at points on a shape, for example, and contrasts attributes (semantic information) from relational (syntactic information) representations. In other places, he discusses statistical and syntactic pattern recognition and builds a contrast between attributes and contextual information. Since the three way relationships between syntax, semantics and statistics were left implicit in his work, we decided to coin Fu's triad—syntax, semantics and statistics—as an explicit homage to his characterization of the

problem space. As one has to be very careful with the word semantics, we now discuss the larger context of the syntax versus semantics problem and explain why interaction is needed to shore up semantics.

In philosophy, semantics is related to the concept of *intentionality*—mental content that is directed to a certain state of the world [24]. For example, certain key points on a shape may be meaningful to us and not easily captured by shape location and normal information. This intentional content directed at shape attributes is meaningful to us and therefore carries semantic information. Intentionality is usually contrasted with *qualia*—the phenomenal qualities of experience [25]. The past twenty years has seen intense debate between those who espouse that machine intentionality will naturally emerge from underlying complex (syntactic) representations and those who believe that intentionality and qualia are coupled and that machines cannot have qualia and consequently, true intentionality as well. Below, we summarize the philosophical state of affairs and discuss the relevance to pattern recognition.

To keep things simple and focused, we—in journalistic mode—first survey the various philosophical positions using a chronological approach. Since the main debate of the past twenty years has been on the ultimate nature of human experience, we begin there. Dennett in [26], set this process in motion by a careful deconstruction of the Cartesian theater—an imagined physical “here and now” in the brain where “it all comes together” and a self is created. Dennett’s deconstruction allowed him to disqualify qualia and adopt an almost eliminativist position regarding human experience. Then, Chalmers in [27], posed the hard problem of consciousness by asking (over and over), why physical and computational processes had to be accompanied by experience. Chalmers argued (convincingly to many) that there was no reason to believe that any process had to be necessarily accompanied by experience and that the processes could very well be “just going on in the dark”. Therefore, there is an *explanatory gap* [28] between physical processes and experience which subsumes the *semantic gap* between linguistic and computational representations. This observation led Chalmers to adopt a dualistic position in which experience is fundamental

and therefore to be added to a physicalist or computational base. The hard problem of experience as formulated by Chalmers was very influential and kick-started a conference series⁴ and an interdisciplinary journal⁵ devoted to the problem. Within this framework, Stoljar, in [29, 30] argued that it is precisely our ignorance of the true nature of the physical that sustains the hard problem. Physicalism could have a mystery ingredient X which always entails experience. Ignorance of this mystery ingredient leads us to adopt dualistic positions prematurely. While the specification of X may be hard, there is no *a priori* reason to rule it out. This position—a fundamental physicalism which entails experience—seemed eminently sound leading us to speculate that *compositionality* is the mystery ingredient [31] and that it is by virtue of being compositional subjects that we have qualia and therefore intentionality as well. Compositional subjects are conceived as physical beings which interact with the world and give it form and are in turn informed. We characterize this position as *non-dualist interactionist* to emphasize the differences with both dualism and eliminativism which preceded it.

We justify the above segue into philosophy by showing the relevance to pattern recognition. An orienting generalization that can be made about the fields of pattern recognition, computer vision and machine learning is their reliance on automated information processing. Insofar as human experience enters the picture, it is merely in the form of behavioral patterns which are modeled—in supervised learning, interactive segmentation, activity recognition etc. While there are exceptions to this stereotyping [32, 33], we feel our generalization is a fair characterization. We argue that this lack is due to the influence of strong AI which is still being felt by these fields long after AI has lost its once dominant position. While Fu’s triad is a constant reminder that syntax and semantics have to be unified, we are not aware of serious attempts at this time to continue along Fu’s path and complete the unification at a fundamental level. With se-

⁴Toward a Science of Consciousness: The Tucson Conference (<http://consciousness.arizona.edu>).

⁵Journal of Consciousness Studies, Imprint Academic, UK.

antics (for us) being related to human experience, we must take Chalmers' point seriously about the need to accommodate experience while eschewing dualism. The nondualist interactionist position outlined above is a step in this direction as it allows for a (physical) subject to interact with a (physical) world with both being informed by the interaction. It has the potential to lead to a unified theory of the behavioral and the intentional while remaining true to its computational roots. The dualism between the physical and experiential is replaced by a naturalistic dualism within physicalism between compositional subjects and processes, thereby evading the hard problem. Since this appears to be a reasonable and common sense approach, we would expect fields such as human-computer interaction (HCI) to have already adopted similar philosophical positions and be presently engaged in constructing a *theory* of HCI by modeling the inter-relationships between the computational, behavioral and intentional worlds.

To the best of our knowledge, no attempt has been made to unify syntax and semantics by attempting to mathematically model (predict and control) the set of interactions between humans and machines while availing of behavioral (activities), intentional (beliefs), computational (objects and patterns) and perhaps even functional (imaging, EEG) data. At present, the field of HCI is more concerned with the design of optimal user interfaces and the ability of interaction to elicit novel phenomenology in test subjects. A focus on the mathematical modeling of human interaction—while well within the purview of HCI—has not yet emerged. This schism between older areas such as pattern recognition and machine learning and newer ones such as HCI and human-centered computing is to be expected since the emphasis on human interaction is very new and clearly a new millennium phenomenon. Obviously there are meetings and conferences—such as the International Conference on Multimodal Interaction (ICMI)—wherein human interaction and pattern recognition come together, but until these become first tier, our point remains valid. Since this trend is expected to continue into the near future, there is a high likelihood that pattern recognition will remain bereft of human interaction (in its broad-

est form as outlined above) while HCI continues exploring novel phenomenology [34]. Clearly, this gap between pattern recognition and HCI mirrors the earlier gap (in the late '70s) between image processing and a fledgling computer vision. Perhaps a postmodern counterpart of Rama is needed to step in and bridge the gap by bringing a sophisticated applied math aesthetic to HCI. When this gap is taken in conjunction with our previous observation that machine learning may continue shedding its syntactic components, we think the loss of Fu will be even more keenly felt in the future.

From the above arguments, we see that there are no principled reasons why pattern recognition cannot include within its rubric the modeling of human interaction with machines. Even if its present day center of narrative gravity is anchored in automated processing, there is no reason to rule out a reboot especially since we witnessed a previous one in the early '80s. We now try to specify one way that things may proceed.

There is an unexpected conceptual connection between interaction as specified here and the wave functions of the previous section. In physics, the only way in which a conscious observer interacts with a physical system is via the process of measurement which causes wave function collapse (from a superposed state) [35]. While this remains a highly controversial subject in physics [36], it dovetails well with a nondualist interactionist position. Since we would like to explicitly take into account different forms of data, we envisage an overall wave function ψ comprising separate components in the form $|\psi_c\rangle|\psi_f\rangle|\psi_i\rangle|\psi_b\rangle$ with the subscripts c , f , i , and b standing for the computational, functional, intentional and behavioral respectively. Here, we have used the standard ket ($|\psi_a\rangle$) notation from physics [37] (loosely understood as an infinite dimensional counterpart of a vector). The notation $|\psi_a\rangle|\psi_b\rangle$ denotes a composite or direct tensor product of kets $|\psi_a\rangle$ and $|\psi_b\rangle$. Of the components above, the intentional presents the most difficult problem for data collection since people can be deluded about their own beliefs—perhaps the main reason why such a unification has not been previously attempted. It may turn out that one way to take this into account is to use separate wave function components for verbal reports and intentions

with the latter being a hidden variable, predicted by the rest. Needless to say, this is all extremely speculative but instructive: there is a clear and pressing need to simultaneously and mathematically model the entire range of human interaction (including behaviors and intentions) in order to unify syntax and semantics. We cannot sketch out a full proposal here but instead attempt to make a limited point about human interaction and shape analysis.

In shape analysis, the wave function discontinuously changes when source locations and/or normals are added/erased. Since wave functions do not evolve in our setup, the discontinuous change is quite natural. It merely corresponds to added/erased source information. Consequently, we may now return to the wave function shape representation and specify a problem of pattern interaction in that space. Note that the issue of creating a model of interaction is orthogonal to the issue of creating a (shape) representation that facilitates interaction. We are merely pointing out that wave functions are felicitous of interaction.

To press this point home, we took a silhouette shape (shown in Figure 4) and obtained a wave function from just a segment of the shape. Curve normal information is not utilized. The wave function is computed using

$$\psi(x) = \sum_{l=1}^L \sum_{k=1}^{N_l} \exp \left\{ -\frac{\|x - x_{kl}\|}{\lambda} \right\} \exp \left\{ i \frac{\|x - x_{kl}\|}{\tau} \right\}, \quad (8)$$

in other words, from location information alone. Shape topology information is extracted from the level sets of the phase as before and clearly show the partialness of each shape segment. However, when we add the wave functions together and display the shape topology from phase, we obtain the full shape figural. We clearly see that this can be operated in an interactive context. A human can interact with this emerging wave function and edit shape locations (and erase them if necessary) to produce the final wave function which contains both topology and location information in phase and magnitude respectively. It remains to be seen if normal information can be included in this framework and the interplay between uncertainty and interaction is poorly understood.

In this section, we took the position that interaction is one of the foundations of intentionality and therefore semantics. In exploring this theme, we

have shown the unexpected relationship between interaction and linearity: when shape source locations are edited—added or erased—the wave function changes discontinuously as expected. But, the linearity of the wave function allows for the discontinuous contributions (in time) to be integrated via linear operations into a new representation. While human interaction in shape wave function representations is new (to our knowledge), it has been repeatedly pointed out in the physics literature. For example, Wheeler in [38] emphasized the centrality of the participant-observer in physics and the importance of the dictum “No elementary phenomenon is a phenomenon until it is a registered (observed) phenomenon”. Since physics is far afield from us, we take our cue from Chalmers who points out [35] in connection with arguments against interactionism: “By far the strongest response to this objection, however, is to suggest that far from ruling out interactionism, contemporary physics is positively encouraging to the possibility.” We have merely taken this point further and in a mathematical direction by modeling interaction of humans with shapes in terms of source location edits leading to discontinuous wave function change—which is entirely straightforward in our context. In the future, we plan to take into account different forms of data (intentional and behavioral) in a wave function framework in the hope of fulfilling a deeper integration of syntax and semantics.

4. Disputation

Knowledge is a deadly friend

When no one sets the rules

King Crimson, Epitaph

Throughout this essay, we have championed Fu’s endeavors in trying to unify syntax, semantics and statistics in pattern recognition. First, we examined his contributions in the light of the differences between syntactic and statistical pattern recognition. The concept of uncertainty was highlighted as a way to soften syntactical representations and make them more robust. When we restricted our focus to shape analysis, we were able to demonstrate the importance of a

new complex wave function whose magnitude and phase simultaneously represented shape location and topology respectively. Next, we looked at the current state of the syntax versus semantics debate and concluded that interaction was a key component missing in present day pattern recognition. An unexpected nexus exists between wave functions and interaction in the physics literature which we leveraged for shape analysis. It turns out that interaction in the form of the editing of source locations and attributes can be accommodated in the wave function representation. Furthermore, properties like linearity and superposition come to our aid when the wave function changes discontinuously in response to interaction. Finally, in the quest for more generality, we suggested that pattern recognition should be open to including human interaction into its fold, perhaps leading to a newer, HIPR (human interaction in pattern recognition) sub-field—a reboot prefigured by the earlier CVPR reconfiguration of the early '80s.

Lest we be accused of being too glib in these recommendations, it's worth noting the gulf that exists between some of the fields under consideration: pattern recognition, human-computer interaction, philosophy and neuroscience. There is no shared consensus between these fields at the present time: for example neuroscience is still focused on how brain processes give rise to experience whereas HCI in the main treats the human as a Cartesian unit. At the same time, philosophy is moving away from *both* the naïve Cartesian dualism of HCI and the materialism of neuroscience. We have already discussed the lack of emphasis on mathematical modeling in HCI in contrast to its centrality in pattern recognition. And pattern recognition, machine learning and computer vision have lately not shown much interest in philosophy preferring to leave intact their tacit commitments to strong AI. Given these divisions, a shotgun synthesis is clearly premature, unwarranted and possibly dangerous. Rather, a synthesis should first be deemed *necessary* perhaps driven by demographic change or the impending dissolution of a field. While the former is to be preferred over the finality of the latter, it too represents a gloaming which is hard to detect until it is almost upon us.

Acknowledgments

We thank David Chalmers, Rama Chellappa and John Corring for helpful conversations.

References

- [1] C. M. Bishop, *Neural networks for pattern recognition*, Oxford University Press, 1996.
- [2] C. M. Bishop, *Pattern recognition and machine learning*, Springer, 2007.
- [3] M. Berman, *All that is solid melts into air: The experience of modernity*, Penguin Books, 1988.
- [4] D. Marr, *Vision: A computational investigation into the human representation and processing of visual information*, W. H. Freeman and Company, New York, NY, 1982.
- [5] K. S. Fu, A step towards unification of syntactic and statistical pattern recognition, *IEEE Trans. Patt. Anal. Mach. Intell.* 8 (3) (1986) 398–404.
- [6] R. Chellappa, R. L. Kashyap, Digital image restoration using spatial interaction models, *IEEE Trans. Acoust., Speech, Signal Processing* 30 (1982) 461–472.
- [7] R. T. Frankot, R. Chellappa, A method for enforcing integrability in shape from shading algorithms, *IEEE Trans. Patt. Anal. Mach. Intell.* 10 (4) (1988) 439–451.
- [8] T. J. Broida, R. Chellappa, Estimation of object motion parameters from noisy images, *IEEE Trans. Patt. Anal. Mach. Intell.* 8 (1) (1986) 90–99.
- [9] B. K. P. Horn, *Robot vision*, MIT Electrical Engineering and Computer Science series, The MIT Press, 1986.

- [10] T. Simchony, R. Chellappa, M. Shao, Direct analytical methods for solving Poisson equations in computer vision problems, *IEEE Trans. Patt. Anal. Mach. Intell.* 12 (5) (1990) 435–446.
- [11] S. Gold, A. Rangarajan, A graduated assignment algorithm for graph matching, *IEEE Trans. Patt. Anal. Mach. Intell.* 18 (4) (1996) 377–388.
- [12] H. Chui, A. Rangarajan, A new algorithm for non-rigid point matching, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 2, IEEE Press, 2000, pp. 44–51.
- [13] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Trans. Patt. Anal. Mach. Intell.* 6 (6) (1984) 721–741.
- [14] H. Putnam, Models and reality, *The Journal of Symbolic Logic* 45 (3) (1980) 464–482.
- [15] J. Searle, Minds, brains and programs, *Behavioral and Brain Sciences* 3 (3) (1980) 417–457.
- [16] S. Russell, P. Norvig, *Artificial intelligence: A modern approach*, 3rd Edition, Prentice Hall, 2009.
- [17] T. Chen, A. Rangarajan, S. J. Eisenschenk, B. C. Vemuri, Construction of a neuroanatomical shape complex atlas from 3D MRI brain structures, *NeuroImage* 60 (3) (2012) 1778–1787.
- [18] S. J. Osher, R. P. Fedkiw, *Level set methods and dynamic implicit surfaces*, Springer, 2002.
- [19] K. S. Gurumoorthy, A. Rangarajan, Distance transform gradient density estimation using the stationary phase approximation, *SIAM J. Math. Analysis* 44 (6) (2012) 4250–4273.

- [20] M. Sethi, A. Rangarajan, K. S. Gurumoorthy, The Schrödinger distance transform (SDT) for point-sets and curves, in: IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 198–205.
- [21] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden variables", I and II, in: J. A. Wheeler, W. H. Zurek (Eds.), Quantum theory and measurement, Princeton Series in Physics, Princeton University Press, 1983, Ch. III.3, pp. 369–402.
- [22] J. Butterfield, On Hamilton-Jacobi theory as a classical root of quantum theory, in: A. Elitzur, S. Dolev, N. Kolenda (Eds.), Quo-vadis quantum mechanics, Springer, 2005, Ch. 13, pp. 239–274.
- [23] M. Born, E. Wolf, Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light, Cambridge University Press, 1999.
- [24] J. Brandl, Intentionality, in: L. Albertazzi, M. Libardi, R. Poli (Eds.), The school of Franz Brentano, Nijhoff International Philosophy Series, Springer, New York, 1996, Ch. 9, pp. 261–264.
- [25] T. Nagel, What's it like to be a bat?, *Philosophical Review* 83 (1974) 435–456.
- [26] D. C. Dennett, *Consciousness explained*, Little, Brown and Company, 1991.
- [27] D. J. Chalmers, *The conscious mind: In search of a fundamental theory*, Oxford University Press, 1996.
- [28] J. Levine, Materialism and qualia: The explanatory gap, *Pacific Philosophical Quarterly* 64 (1983) 354–361.
- [29] D. Stoljar, *Ignorance and imagination: The epistemic origin of the problem of consciousness*, Oxford University Press, 2006.
- [30] D. Stoljar, *Physicalism (New Problems in Philosophy)*, Routledge, 2010.

- [31] A. Rangarajan, Panpsychism, emergence and physicalism, in: *Toward a Science of Consciousness*, Tucson, AZ, 2012.
- [32] G. Nagy, J. Zou, Interactive visual pattern recognition, in: *International Conference on Pattern Recognition (ICPR)*, Vol. 2, 2002, pp. 478–481.
- [33] E. Vidal, L. Rodríguez, F. Casacuberta, I. García-Varea, Interactive pattern recognition, in: *Fourth Intl. Conference on Machine Learning for Multimodal Interaction (MLMI)*, 2007, pp. 60–71.
- [34] S. Benford, C. Greenhalgh, G. Giannachi, B. Walker, J. Marshall, T. Rodden, Uncomfortable interactions, in: *Conference on Human factors in Computing Systems (CHI)*, 2012, pp. 2005–2014, best paper.
- [35] D. J. Chalmers, Consciousness and its place in nature, in: D. J. Chalmers (Ed.), *Philosophy of mind: Classical and contemporary readings*, Oxford University Press, 2002, Ch. 2:A-27, pp. 247–272.
- [36] D. Z. Albert, *Quantum mechanics and experience*, Harvard University Press, Cambridge, MA, 1994.
- [37] P. A. M. Dirac, A new notation for quantum mechanics, *Mathematical Proceedings of the Cambridge Philosophical Society* 35 (3) (1939) 416–418.
- [38] J. A. Wheeler, Law without law, in: J. A. Wheeler, W. H. Zurek (Eds.), *Quantum theory and measurement*, Princeton Series in Physics, Princeton University Press, 1983, Ch. I.13, pp. 182–216.

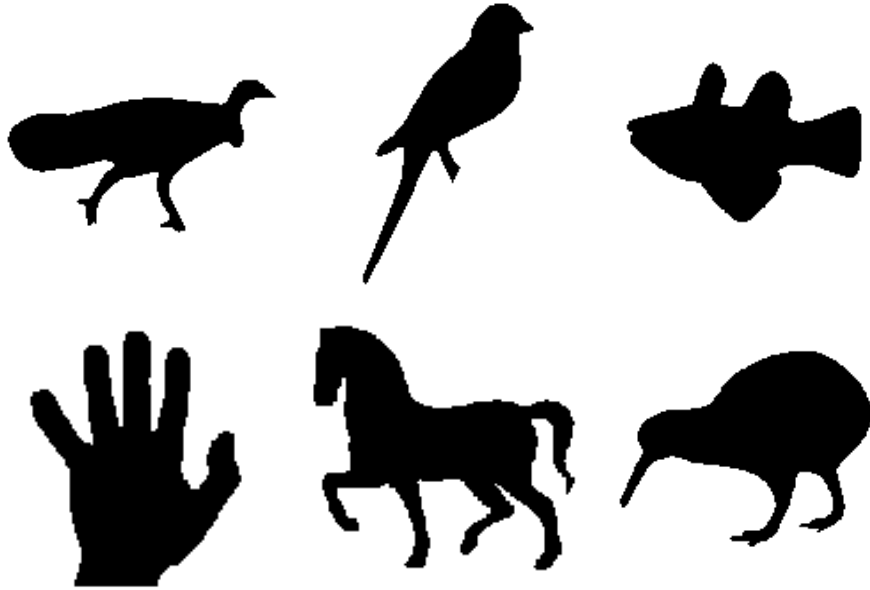


Figure 1: Six silhouette shapes.

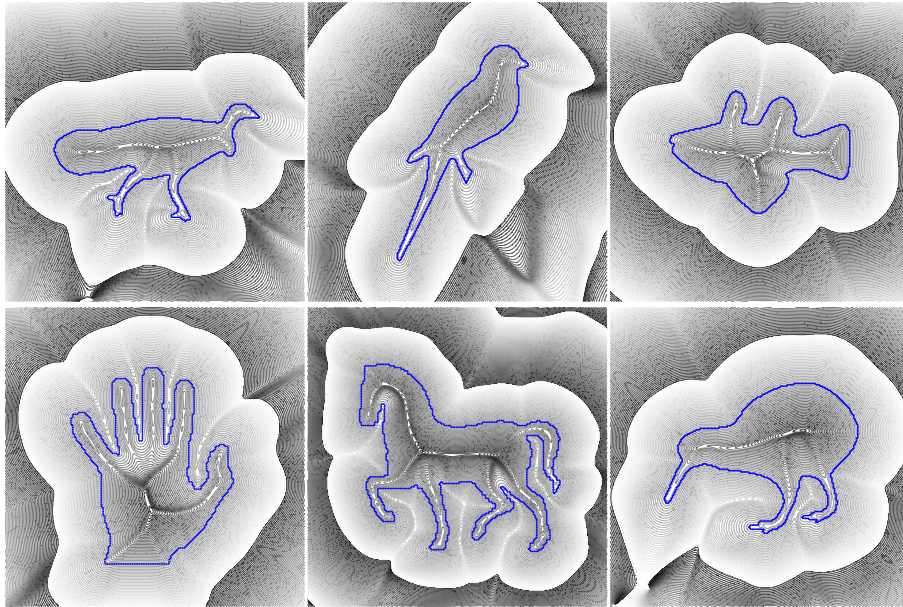


Figure 2: Shape topology information from phase for six silhouettes. Please zoom into the online version for a more detailed view.

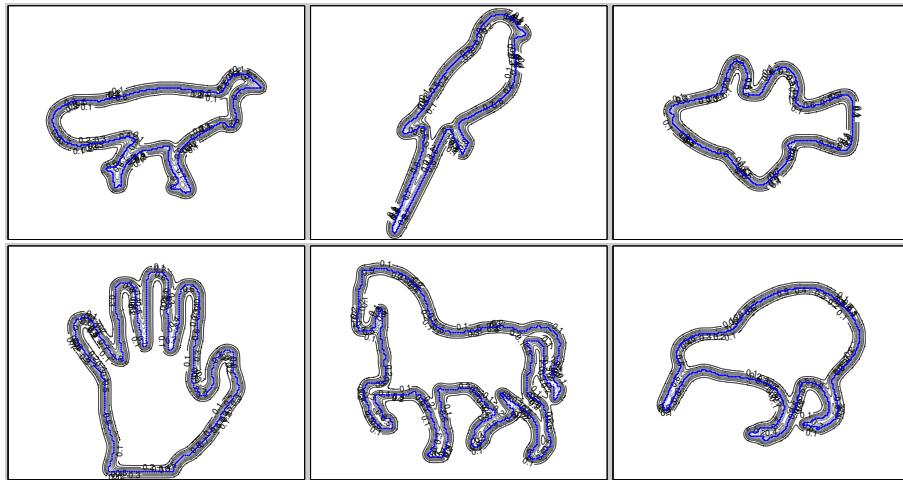


Figure 3: Shape location information from magnitude for six silhouettes. Please zoom into the online version for a more detailed view.

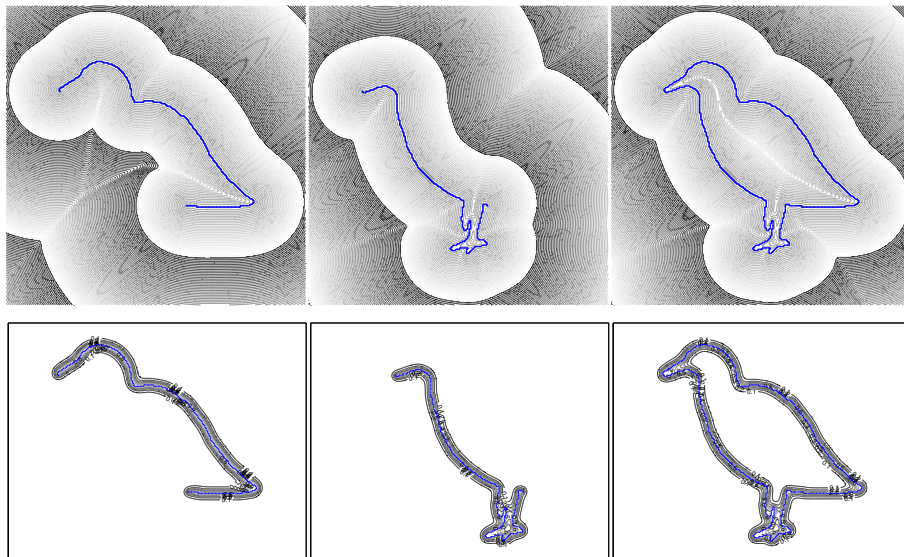


Figure 4: Shape addition. Top: Bird image. Second Row: Left and center: shape topologies from phase for partial segments. Right: shape topology obtained from the sum of the wave functions. Bottom Row: Left and center: shape location from magnitude for partial segments. Right: shape location from magnitude obtained from the sum of the wave functions. Please zoom into the online version for a more detailed view.