Mirror Neurons, Mirrorhouses, and the Algebraic Structure of the Self

Ben Goertzel[1], Onar Aam, F. Tony Smith, Kent Palmer
[1]Novamente LLC

November 18, 2007

Abstract. Recent psychological research suggests that the individual human mind may be effectively modeled as involving a group of interacting social actors: both various subselves representing coherent aspects of personality; and virtual actors embodying "internalizations of others." Recent neuroscience research suggests the further hypothesis that these internal actors may in many cases be neurologically associated with collections of mirror neurons. Taking up this theme, we study the mathematical and conceptual structure of sets of inter-observing actors, noting that this structure is mathematically isomorphic to the structure of physical entities called "mirrorhouses." Mirrorhouses are naturally modeled in terms of abstract algebras such as quaternions and octonions (which also play a central role in physics), which leads to the conclusion that the presence within a single human mind of multiple inter-observing actors naturally gives rise to a mirrorhouse-type cognitive structure and hence to a quaternionic and octonionic algebraic structure as a significant aspect of human intelligence. Similar conclusions would apply to nonhuman intelligences such as AI's, we suggest, so long as these intelligences included empathic social modeling (and/or other cognitive dynamics leading to the creation of simultaneously active subselves or other internal autonomous actors) as a significant component.

Introduction

The thesis of this paper is that there are certain abstract algebraic structures that typify the self-structure of human beings and any other intelligent systems relying on empathy for social intelligence. These algebraic structures, called quaternions and octonions, are familiar to mathematicians, and also play a critical role in modern theoretical physics (Dixon, 1994).

The argument presented in favor of this thesis has two steps. First, it is argued that much of human psychodynamics consists of "internal dialogue" between separate internal actors – some of which may be conceived as subselves a la (Rowan, 1990), some of which may be "virtual others" intended to explicitly mirror other humans (or potentially other entities like animals or software programs). Second, it is argued that the structure of inter-observation among multiple inter-observing actors naturally leads to quaternionic and octonionic algebras. Specifically, the structure of inter-observation among three inter-observers is quaternionic; and the structure of inter-observation among four inter-observers is octonionic. This mapping between inter-observation and abstract algebra is made particularly vivid by the realization that the quaternions model the physical situation of three mirrors facing each other in a triangle; whereas the octonions model the physical situation of four mirrors facing each other in a tetrahedron, or more complex packing structures related to tetrahedra. Using these facts, we may phrase the main thesis of the current paper in a simple form: ***The structure of the self of an empathic social intelligence is that of a quaternionic or octonionic mirrorhouse.***

There are obvious echoes here of Buckminster Fuller's (1982) philosophy, which viewed the tetrahedron as an essential structure for internal and external reality. And there is a next step that is even more Fulleresque in nature: the structure of a group of socially interacting individu als is that of a tiling of part of space using adjacent quaternionic or octonionic mirrorhouses.

There is also an intriguing potential tie-in with recent developments in neurobiology, which suggest that empathic modeling of other minds may be carried out in part via a "mirror neuron system" that enables a mind to experience another's actions, in a sense, "as if they were its own" (Ramachandran, 2006). Building on existing data and theories regarding mirror neurons, we hypothesize a "neural mirrorhouse system" supporting a mirrorhouse-structured self.

In the remainder of the paper, we will explore these issues from the perspective of psychology, biology and mathematics, finally returning at the end to a discussion of what it means phenomenologically, from the internal perspective of the experiencing mind.

The Intrinsic Sociality of the Self

In what sense may it be said that the self of an individual human being is a "social" system?

A special issue of "Journal of Consciousness Studies" (Thompson, 2001) provides an excellent summary of recent research and thinking on this topic. A basic theme of several of the papers is as follows:

- The human brain contains structures specifically configured to respond to other humans' behaviors (these appear to involve "mirror neurons" and associated "mirror neuron systems," on which we will elaborate below).
- these structures are also used internally when no other people (or other agents) are present, because human self is founded on a process of continual interaction between "phenomenal self" and "virtual other(s)", where the virtual others are reflected by the same neural processes used to mirror actual others
- so, the iteration between phenomenal self and actual others is highly wrapped up with the interaction between phenomenal self and virtual others

This line of research focuses on exploration of the dynamics by which self is fundamentally grounded in sociality and social interactions. The social interactions that structure the self are in part grounded in the interactions between the brain structures generating the phenomenal self and the brain structures generating the virtual others. That is, they are part of the dynamics of the self as well as part of the interactions between self and actual others. The key point is that human self is intrinsically not autonomous and independent, but rather is intrinsically dialogic and intersubjective.

Another way to phrase this is in terms of "empathy." That is, one can imagine an intelligence that attempted to understand other minds in a purely impersonal way, simply by reasoning about their behavior. But that doesn't seem to be the whole story of how humans do it. Rather, we do it, in part, by running simulations of the other minds internally – by spawning virtual actors, virtual selves within our own minds that emulate these other actors (according our own understanding). This is why we have the feeling of empathy – of feeling what another mind is feeling. It's because we actually *are* feeling what the other mind is feeling – in an approximation, because we're feeling what our internal simulation of the other mind is feeling. Thus, one way to define "empathy" is as the understanding of other minds via internal simulation of them. Clearly, internal simulation is not the only strategy the human mind takes to studying other minds – the patterns of errors we make in predicting others' behaviors indicates that there is also an inferential, analytical component to a human's understanding of others (Carruthers and Smith, 1996). But empathic simulation is a key component, and we suggest that, in normal humans (autistic humans may be a counterexample; see Oberman et al, 2005), it is the most central aspect of other-modeling, the framework upon which other sorts of other-modeling such as inferencing are layered.

This perspective has some overlap with John Rowan's (1990) theory of human subpersonalities, according to which each person is analyzed as possessing multiple subselves representing different aspects of their nature appropriate to different situations. Subselves may possess different capabilities, sometimes different memories, and commonly differently biased views of the common memory store. Numerous references to this sort of "internal community" of the mind exist in literature, e.g. Proust's reference to "the several gentlemen of whom I consist."

Putting these various insights together, we arrive at a view of the interior of the human mind as consisting of not a single self but a handful of actors representing subselves and virtual others. In other words, we arrive at a perspective of human mind as social mind, not only in the sense that humans define themselves largely in terms of their interactions with others, but also in the sense that humans are substantially internally constituted by collections of interacting actors each with some level of self-understanding and autonomy. The primary contribution of this paper is a specific hypothesis regarding the structure of this internal social mind: that is corresponds to the structures of certain physical constructs (mirrorhouses) and certain abstract algebras (quaternions, octonions and Clifford algebras).

Mirror Neurons and Associated Neural Systems

A number of thinkers have tied together the above ideas regarding self-as-social-system, with recent neurobiological results regarding the role of mirror neurons and associated neural systems in allowing human and animal minds to interpret, predict and empathize with other human and animal minds with which they interact. The biology of mirror neuron systems is still only partially understood, so that the tie-in between mirror neurons and psychological structures posited here must be viewed as subject to revision based on further refinement of our understanding in the biology of mirror neurons. Ultimately, the core ideas of this paper would remain equally valid if one replaced "mirror neurons" and associated systems with some other, functionally similar neural mechanism. However, given that we do have some reasonably solid biological data -- and some additional, associated detailed biological hypotheses – regarding the role of mirror neurons in supporting the functions of empathy and self, it is interesting to investigate what these data and hypotheses suggest.

In simplest terms, a mirror neuron is a neuron which fires both when an animal acts and when the animal observes the same action performed by another animal, especially one of the same species. Thus, the neuron is said to "mirror" the behavior of another animal – creating a similar neuronal activation patterns as if the observer itself were acting. Mirror neurons have been directly observed in primates, and are believed to exist in humans as well as in some other mammals and birds (Blakeslee, 2006). Evidence suggestive of mirror neuron activity has been found in human premotor cortex and inferior parietal cortex. V.S. Ramachandran (2006) has been among the more vocal advocates of the important of mirror neurons, arguing that they may be one of the most important findings of neuroscience in the last decade, based on the likelihood of their playing a strong role in language acquisition via imitative learning.

The specific conditions under which mirror neuron activity occurs are still being investigated and are not fully understood. Among the classic examples probed in lab experiments are grasping behavior, and facial expressions indicating emotions such as disgust. When an ape sees another ape grasp something, or make a face indicating disgust, mirror neurons fire in the observing ape's brain, similar to what would happen if the observing ape were the one doing the grabbing or experiencing the disgust. This is a pretty powerful set of facts – what it says is that shared experience among differently embodied minds is not a purely cultural or psychological phenomenon, it's something that is wired into our physiology. We really can feel each others' feelings as if they were our own; to an extent, we may even be able to will each others' actions as if they were our own (Lohmar, 2006).

Equally interesting is that mirror neuron response often has to do with the perceived intention or goal of an action, rather than the specific physical action observed. If another animal is observed carrying out an action that is expected to lead to a certain goal, the observing animal may experience neural activity that it would experience if it had achieved this goal. Furthermore, mere visual observation of actions doesn't necessarily elicit mirror neuron activity. Recent studies (Buccino et al, 2001, 2004) involved scanning the brains of various human subjects while they were observing various events, such another person speaking or biting something, a monkey lip-smacking or a dog barking. The mirror neurons were not activated by the sight of the barking dog – presumably because this was understood visually and not empathically (since people don't bark), but were activated by the sight of other people as well as of monkeys.

There is also evidence that mirror neurons may come to be associated with learned rather than just inherited capabilities. For instance, monkeys have mirror neurons corresponding to specific activities such as tearing paper, which are learned in the lab and have no close correlate in the wild (Rizzolatti, 2004).

Perhaps the most ambitious hypothesis regarding the role of mirror neurons in cognition is Rizzolatti and Arbib's (1998) Mirror System Hypothesis, which conjectures that neural assemblies reliant on mirror neurons played a key role in the evolution of language. These authors suggest that Broca's area (associated with speech production) evolved on top of a mirror system specialized for grasping, and inherited from this mirror system a robust capacity for pattern recognition and generation, which was then used to enable imitation of vocalizations, and to encourage "parity" in which associations involving vocalizations are roughly the same for the speaker as for the hearer. According to the MSH, the evolution of language proceeded according to the following series of steps (Arbib et al, 2006):

- S1: Grasping.
- S2: A mirror system for grasping, shared with the common ancestor of human and monkey.
- S3: A system for simple imitation of grasping shared with the common ancestor of human and chimpanzee. The next 3 stages distinguish the hominid line from that of the great apes:
- S4: A complex imitation system for grasping.
- S5: Protosign, a manual-based communication system that involves the breakthrough from employing manual actions for praxis to using them for pantomime (not just of manual actions), and then going beyond pantomime to add conventionalized gestures that can disambiguate pantomimes.
- S6: Protospeech, resulting from linking the mechanisms for mediating the semantics of protosign to a vocal apparatus of increasing flexibility. The hypothesis is not that S5 was completed before the inception of S6, but rather that protosign and protospeech evolved together in an expanding spiral.
- S7: Language: the change from action-object frames to verb-argument structures to syntax and semantics.

As will be discussed below, one may correlate this series of stages with a series of mirrorhouses involving an increasing number of mirrors. This leads to an elaboration of the MSH, which posits that evolutionarily, as the mirrorhouse of self and attention gained more mirrors, the capability for linguistic interaction became progressively more complex.

Quaternions and Octonions

In this section, as a preparation for our mathematical treatment of mirrorhouses and the self, we review the basics of the quaternion and octonion algebras. This is not original material, but it is repeated here because it is not well known outside the mathematics and physics community. Readers who want to learn more should follow the references.

Most readers will be aware of the real numbers and the complex numbers. The complex numbers are formed by positing an "imaginary number" i so that i*i=-1, and then looking at "complex numbers" of the form a+bi, where a and b are real numbers. What is less well known is that this approach to extending the real number system may be generalized further. The quaternions are formed by positing three imaginary numbers i, j and k with i*i=j*j=k*k=-1, and then looking at "quaternionic numbers" of the form a + bi + cj + dk. The octonions are formed similarly, by positing 7 imaginary numbers i,j,k,E,I,J,K and looking at "octonionic numbers" defined as linear combinations thereof.

Why 3 and 7? This is where the math gets interesting. The trick is that only for these dimensionalities can one define a multiplication table for the multiple imaginaries so that unique division and length measurement (norming) will work. For quaternions, the "magic multiplication table" looks like

i*j = k	j*i = -k
j*k = i	k*j = -i
k*i = j	i*k = -j

Using this multiplication table, for any two quaternionic numbers A and B, the equation

$$x * A = B$$

has a unique solution when solved for x. Quaternions are not commutative under multiplication, unlike real and complex numbers: this can be seen from the above multiplication table in which e.g. i*j is not equal to j*i. However, quaternions are normed: one can define ||A|| for a quaternion A, in the familiar root-mean-square manner, and get a valid measure of length fulfilling the mathematical axioms for a norm.

Note that you can also define an opposite multiplication for quaternions: from i*j = k you can reverse to get j*i = k, which is an opposite multiplication, that still works, and basically just constitutes a relabeling of the quaternions. This is different from the complex numbers, where there is only one workable way to define multiplication.

The quaternion algebra is fairly well known due to its uses in classical physics and computer graphics (Hanson, 2006); the octonion algebra, also known as Cayley's octaves, is less well known but is adeptly reviewed by John Baez (2002).

The magic multiplication table for 7 imaginaries that leads to the properties of unique division and normed-ness is as follows:

1	**i**	**j**	**k**	**E**	**I**	**J**	**K**
i	-1	k	-j	I	-E	-K	J
j	-k	-1	i	J	K	-E	-I
k	j	-i	-1	K	-J	I	-E
E	-I	-J	-K	-1	i	j	k
I	E	-K	J	-i	-1	-k	j
J	K	E	-I	-j	k	-1	-i
K	-J	I	E	-k	-j	i	-1

Actually this is just one of 480 basically equivalent (and equally "magical") forms of the octonionic multiplication table (as opposed to the 2 varieties for quaternions, mentioned above). Note that, according to this or any of the other 479 tables, octonionic multiplication is neither commutative nor associative; but octonions do satisfy a weaker form of associativity called alternativity, which means that the subalgebra generated by any two elements is associative.

As it happens, the only normed division algebras over the reals are the real, complex, quaternionic and octonionic number systems. These four algebras also form the only alternative, finite-dimensional division algebras over the reals. These theorems are nontrivial to prove, and fascinating to contemplate.

Modeling Mirrorhouses Using Quaternions and Octonions

Now let's move from algebras to mirrors – houses of mirrors, to be precise. Interestingly shaped houses of mirrors!

Mirrorhouses are structures built up from mutually facing mirrors which reflect each others' reflections. The simplest mirrorhouse possible to construct is made of two facing mirrors, X and Y. X reflects Y and Y reflects X.

One convenient way to describe mirrorhouses is to introduce hypersets, as described e.g. in (Barwise and Etchemendy, 1989). Hypersets are mathematical sets that are freed from the Axiom of Foundation, so that unlike an ordinary mathematical set, a hyperset A may contain A as an element, or may contain a set that contains A as an element, etc. The general utility of hypersets for modeling complex systems is discussed in (Goertzel, 1994).

In terms of hypersets, a simple 2-mirror mirrorhouse may be crudely described as:

$$X = \{Y\}$$

$$Y = \{X\}$$

(ignoring the inversion effect of mirroring).

Note that if we try to unravel this hyperset by inserting one element into the other we arrive at an infinite regress:

$$Y = \{X = \{Y = \{X = \{Y = \{X = \{Y = \{\{X = \{Y = \{...\} \} \} \} \} \} \} \} \}$$

This corresponds to the illusory infinite tube which interpenetrates both mirrors.

Suppose now that we constructed a mirrorhouse from *three* mirrors instead of two. What hyper-structure would this have? Amazingly it turns out that it has precisely the structure of the quaternion imaginaries.

Let i, j and k be hypersets representing three facing mirrors. We then have that

$$i = \{j,k\}$$

$$j = \{k,i\}$$

and

$$k = \{i,j\}$$

where the notation $i=\{j,k\}$ means, e.g. that mirror i reflects mirrors j and k in that order.

With three mirrors ordering now starts playing a vital role because mirroring inverts left/right-handedness. If we denote the mirror inversion operation by "-" we have that

$$i = \{j,k\} = -\{k,j\}$$

$$j = \{k,i\} = -\{i,k\}$$

and

$$k = \{i,j\} = -\{j,i\}$$

But the above is exactly the structure of the quaternion triple of imaginaries:

$$i = j*k = -k*j$$

$$j = k*i = -i*k$$

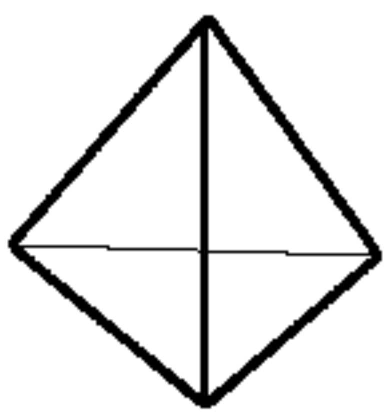
$$k = i*j = -j*i$$

The quaternion algebra therefore is the precise model of the holographic hyper-structure of three facing mirrors, where we see mirror inversion as the quaternionic anti-commutation. The two versions of the quaternion multiplication table correspond to the two possible ways of arranging three mirrors into a triangular mirrorhouse.

When we move on to octonions, things get considerably subtler – though no less elegant, and no less conceptually satisfying. While there are 2 possible quaternionic mirrorhouses, there are 480 possible octonionic mirrorhouses, corresponding to the 480 possible variant octonion multiplication tables!

Recall that the octonions have 7 imaginaries i,j,k,E,I,J,K, which have 3 algebraic generators i,j,E (meaning that combining these three imaginaries can give rise to all the others). The third generator E is distinguished from the others, and we can vary it to get the 480 multiplications/mirrorhouses.

The simplest octonionic mirrorhouse is simply the tetrahedron:



More complex octonionic mirrorhouses correspond to tetrahedra with extra mirrors placed over their internal corners. This gives rise to very interesting geometric structures, which have been explored by Buckminster Fuller and also by various others throughout history.

Start with a 3-dimensional tetrahedron of 4 facing mirrors. Let the floor be the distinguished third generator E and the 3 walls be I,J,K (with a specific assignment of walls to imaginaries, of course). Then, by reflection through the E floor, the reflected I J K become i j k, and we now have all 7 imaginary octonions. This relatively simple tetrahedral mirrorhouse corresponds to one of the 480 different multiplications; the one given in the table above.

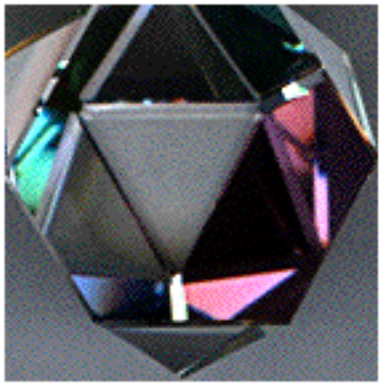
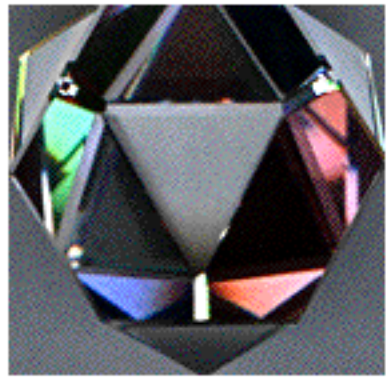
To get another we truncate the tetrahedron. Truncation puts a mirror parallel to the floor, making a mirror roof. Then, when you look up at the mirror roof, you see the triangle roof parallel to the floor E. The triangle roof parallel to the floor E represents the octonion -E, and reflection in the roof -E gives 7 imaginary octonions with the multiplication rule in which -E is the distinguished third generator.

Looking up from the floor, you will also see 3 new triangles having a common side with the triangle roof -E, and 6 new triangles having a common vertex with the triangle roof -E.

The triangle roof + 9 triangles = 10 triangles form half of the faces (one hemisphere) of a 20-face quasi-icosahedron. The quasi-icosahedron is only qualitatively an icosahedron, and is not exact, since the internal angle of the pentagonal vertex figure of the reflected quasi-icosahedron is not 108 degrees, but is 109.47 degrees (the octahedral dihedral angle), and the vertex angle is not 72 degrees, but is 70.53 degrees (the tetrahedral dihedral angle). (To get an exact icosahedral kaleidoscope, three of the triangles of the tetrahedron should be golden isosceles triangles.)

Each of the 9 new triangles is a "reflection roof" defining another multiplication. Now, look down at the floor E to see 9 new triangles reflected from the 9 triangles adjoining the roof -E. Each of these 9 new triangles is a "reflection floor" defining another multiplication. We have now 1 + 1 + 9 + 9 = 20 of the 480 multiplications.

Just as we put a roof parallel to the floor E by truncating the top of the tetrahedral pyramid, we can put in 3 walls parallel to each of the 3 walls I, J, K by truncating the other 3 points of the tetrahedron, thus getting 3x20 = 60 more multiplications. That gives us 20 + 60 = 80 of the 480 multiplications.

To get the rest, recall that we fixed the walls I, J, K in a particular order with respcet to the floor E. There are 3! = 6 permutations of the walls I, J, K Taking them into account, we get all 6x80 = 480 multiplications.

In mathematical terms, this approach effectively fixes the 20-face quasi-icosahedron and varies the 4 faces of the EIJK tetrahedron according to the 24-element binary tetrahedral group {3,3,2} = SL(2,3) to get the 20x24 = 480 multiplications.

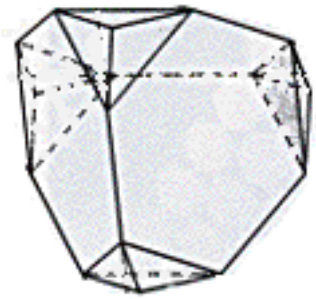
Note that the truncated tetrahedron with a quasi-icosahedron at each vertex combines two types of symmetries:

- **tetrahedral**, related to the square and the ratio sqrt(2), which gives open systems like: an arithmetic series overtone acoustic musical scale with common difference 1/8; the Roman Sacred Cut in architecture; and multilayer space-filling cuboctahedral crystal growth.
- **icosahedral**, related to the pentagon, the Golden Mean (aka Golden Section), and Fibonacci sequences, which gives closed systems like: a harmonic pentatonic musical scale; Le Corbusier's Modulor; and single-layer icosahedral crystals.

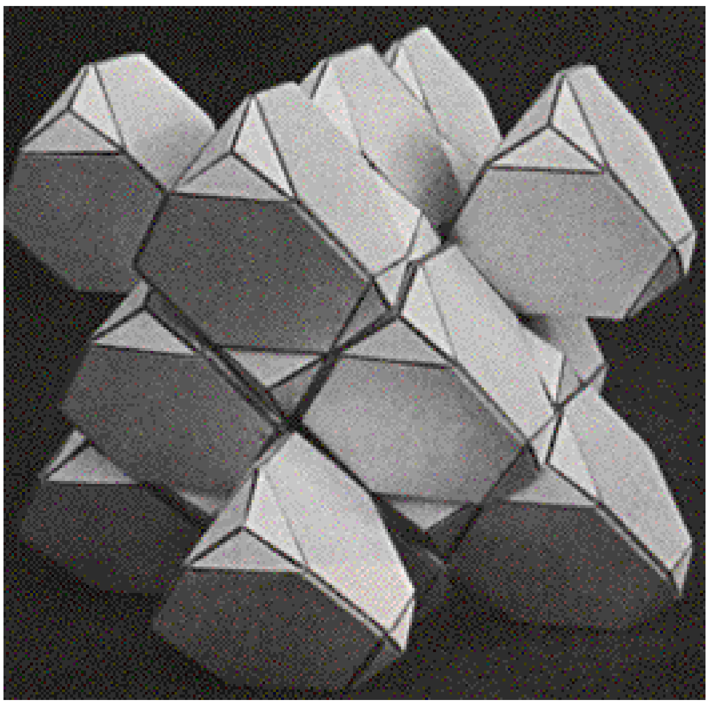
It is interesting to observe that the binary icosahedral group is isomorphic to the binary symmetry group of the 4-simplex, which may be called the pentahedron and which David Finkelstein and Ernesto Rodriguez (1984) have called the “Quantum Pentacle.” A pentahedron has 5 vertices, 10 edges, 10 areas, and 5 cells. The 10 areas of a pentahedron correspond to the 10 area faces of one hemisphere of an icosahedron.

The pentahedron projected into 3 dimensions looks like a tetrahedron divided into 4 quarter-tetrahedra.

If you add a quarter-tetrahedron to each truncation of a truncated tetrahedron,



you get a space-filling polytope that can be centered on a vertex of a 3-dimensional diamond packing to form a Dirichlet domain of the 3-dimensional diamond packing.



(A Dirichlet domain of a vertex in a packing is the set of points in the space in which the packing is embedded that are nearer to the given vertex than to any other.) The 4 most distant vertices of the Dirichlet domain polytope are vertices of the dual diamond packing in 3-dimensional space.

All in all, we conclude that:

- In its simplest form the octonion mirrorhouse is a tetrahedral mirrorhouse
- In its more general form, the octonion mirrorhouse shows a tetrahedral diamond packing network of quasi-icosahedra, or equivalently, of quasi-pentahedra

Observation as Mirroring

Now we proceed to draw together the threads of the previous sections: mirror neurons and subselves, mirrorhouses and normed division algebras.

To map the community of actors inside an individual self into the mirrorhouse/algebraic framework of the previous section, it suffices to interpret the above

X = {Y}
Y = {X}

as

"X observes Y"
"Y observes X"

(e.g. we may have X= primary subself, Y=inner virtual other), and the above

i = {j,k}
j = {k,i}
k = {i,j}

as

"i observes {j observing k}"
"j observes {k observing i}"
"k observes {i observing j}"

Then we can define the - observation as an inverter of observer and observed, so that e.g.

{j,k} = -{k,j}

We then obtain the quaternions

i = j*k = -k*j
j = k*i = -i*k
k = i*j = -j*i

where multiplication is observation and negation is reversal of the order of observation. Three inter-observers = quaternions.

The next step is mathematically natural: if there are four symmetric inter-observers, one obtains the octonions, according to the logic of the above-described tetrahedral/tetrahedral-diamond-packing mirrorhouse. Octonions may also be used to model various situations involving more than four observes with particular asymmetries among the observers (the additional observers are the corner-mirrors truncating the tetrahedron.)

Why not go further? Who's to say that the internal structure of a social mind isn't related to mirrorhouses obtained from more complex shapes than tetrahedra and truncated tetrahedra? This is indeed not impossible, but intuitively, we venture the hypothesis that where human psychology is concerned, the octonionic structure is complex enough. Going beyond this level one loses the normed division-algebra structure that makes the octonions a reasonably nice algebra, and one also gets into a domain of dramatically escalated combinatorial complexity.

Biologically, what this suggests is that the MSH of Rizzolatti and Arbib just scratches the surface. The system of mirror neurons in the human mind may in fact be a "mirrorhouse system," involving four different cell assemblies, each involving substantial numbers of mirror neurons, and arranged in such a manner as to recursively reflect and model one another. This is a concrete neurological hypothesis which is neither strongly suggested nor in any way refuted by available biological data: the experimental tools at our current disposal are simply not adequate to allow empirical exploration of this sort of hypothesis. The empirical investigation of cell assembly activity is possible now only in a very primitive way, using crude tools such as voltage-sensitive dyes which provide data with a very high noise level (see e.g. Collins et al, 2007). Fortunately though, the accuracy of neural measurement technology is increasing at an exponential rate (Kurzweil, 2005), so there is reason to believe that within a few decades hypotheses such as the presently positive "neural mirrorhouse" will reside in the domain of concretely-explorable rather than primarily-theoretical science.

And finally, we may take this conceptual vision one more natural step. The mirrorhouse inside an individual person's mind is just one small portion of the overall social mirrorworld. What we really have is a collection of interlocking mirrorhouses. If one face of the tetrahedron comprising my internal mirrorhouse at a certain moment corresponds to one of your currently active subselves, then we may view our two selves at that moment as two adjacent tetrahedra. We thus arrive at a view of a community of interacting individuals as a tiling of part of space using tetrahedra, a vision that would have pleased Buckminster Fuller very much indeed.

Phenomenology of Mirroring

But what does all this abstraction mean in terms of individual subjective experience?

Lohmar (2006) has explored the experiential significance of mirror neurons using the language of phenomenology. In this vein he has proposed several theses:

- **Thesis 1: Maximality.** We can co-experience all dimensions of experiencing in other persons.

- **Thesis 2: Weakness**. In co-experiencing the experiences of other persons we always deal with an experience that is dimmed or weakened in a characteristic way.
- **Thesis 3: Phantasmata.** Co-experienced sensations are "phantasma" of sensations. A phantasma of a sensation is "something like" a sensation, i.e., it is given to us in the medium of a sensation; but it is not, however, a real sensation, because phantasmata take place in the absence of that which normally evokes the appropriate sensation. The phantasmata, which make our co-sensing possible, do not appear deliberately but rather unwillingly. But the fact that they occur unwillingly does not imply that they occur automatically in all cases. In Husserlian language, it may be said that phantasmata have both a sense-bearing and sense-fulfilling function at the same time (Husserl 1970, Section 9).
- **Thesis 4: Sense-bearing intentions**. Phantasmata with which we co-experience the sensations, feelings, volition and bodily actions of others have a precise sense. They are specific intentions-of-something, i.e., they are sense-bearing intentions.

Lohmar rephrases his fourth thesis in terms of the idea of "co-willing with others" – the idea that we may experience the doing of something when someone else does it. While this may appear problematic, the issues go away when one delves into the neuropsychology of experienced "free will," which is well-documented to largely consist of post-facto explanations of unconsciously-determined actions (Freeman et al, 2000). If ordinary cases of will are largely "illusory" in this sense, there is no reason why instances of co-willing can't have the same phenomenological and neurophysiological status as instances of individual willing.

In Lohmar's terminology, we may say ***that the various observers inside an individual's mental mirrorhouse are recursively experiencing each others' phantasmal sensations as higher-order phantasmata*** – we have phantasmata of phantasmata of phantasmata ... and sometimes there is a real sensation in there too, getting reflected around and around; but there need not necessarily be. One may also have phantasmata that merely reflect other phantasmata, in a bottomless non-well-founded hierarchy. This reminds one of Baudrillard's (1983) notion of simulation, as a process that in itself need not be simulating anything real. One of the purposes of reflection is simulation; and one of the main uses of simulation is to simulate physically real phenomena; but this is not the only possible use.

Specific Instances of Mental Mirrorhousing

What does all this mean in terms of our everyday lives?

Most examples of mental mirrorhousing, I suggest, are difficult for us to distinguish introspectively from other aspects of our inner lives. Mirroring among multiple subselves, simulations of others and so forth is so fully woven into our consciousness that we don’t readily distinguish it from the rest of our inner life. Because of this, the nature of mental mirroring is most easily understood via reference to “extreme cases.”

For instance, consider the following rather mundane real-life situation: Ben needs to estimate the time-duration of a software project that has been proposed for the consulting division of his AI software company. Ben knows he typically underestimates the amount of time required for a project, but that he can usually arrive at a more accurate estimate via conversation with his colleague Cassio. But Cassio isn't available at the moment; or Ben doesn’t want to bother him. So, Ben simulates an "internal Cassio," and they dialogue together, inside Ben’s “mind’s eye.” This is a mirror facing a mirror – an internal Ben mirroring an internal Cassio.

But this process in itself may be more or less effective depending on the specifics -- depending on, for example, which aspects of Ben or Cassio are simulated. So, an additional internal observing mind may be useful for, effectively, observing multiple runs of the "Ben and Cassio conversation simulator" and studying and tuning the behavior. Now we have a quaternionic mirrorhouse.

But is there a deeper inner observer watching over all this? In this case we have an octonionic, tetrahedral mirrorhouse.

The above is a particularly explicit example -- but we suggest that much of everyday life experience consists of similar phenomena, where the different inter-mirroring agents are not necessarily associated with particular names or external physical agents, and thus are more difficult to tangibly discussed. As noted above, this relates closely to Rowan’s analysis of human personality as consisting largely of the interactional dynamics of various never-explicitly-articulated and usually-not-fully-distinct subpersonalities.

For another sort of example, consider the act of creativity, which in (Goertzel, 1997) is modeled in terms of a "creative subself": a portion of the mind that is specifically devoted to creative activity in one more more media, and has its own life and awareness and memory apart from the primary self-structure. The creative subself may create a work, and present it to the main subself for consideration. The three of these participants -- the primary subself, the creative subself and the creative work -- may stand in a relationship of quaternionic mirroring. And then the meta-self who observes this threefold interaction completes the tetrahedral mirrorhouse.

Next, let us briefly consider the classic Freudian model of personality and motivation. According to Freud (1962), much of our psychology consists of interaction between ego, superego and id. Rather than seeking to map the precise Freudian notions into the present framework, we will briefly comment on how ideas inspired by these Freudian notions might play a role in the present framework. The basic idea is that, to the extent that there are neuropsychological subsystems corresponding to Freudian ego, superego and id, these subsystems may be viewed as agents that mirror each other, and hence as a totality may be viewed as a quaternionic mirrorhouse. More specifically we may correlate

- **ego** with the neuropsychological structure that Thomas Metzinger (2004) has identified as the "phenomenal self"
- **superego** with the neuropsychological structure that represents the mind's *learned goal system* -- the set of goals that the system has created
- **id** with the neuropsychological structure that represents the mind's *in-built goal system*, which largely consists of basic biological drives

Using this interpretation, we find that a quaternionic ego/superego/id mirrorhouse may indeed play a role in human psychology and cognition. However, there is nothing in the theoretical framework being pursued here to suggest that this particular configuration of inter-observers has the foundational significance Freud ascribed to it. Rather, from the present perspective, this Freudian triarchy appears as important configuration (but not the only one) that may arise within the mirrorhouse of focused attention.

And, of course, if we add in the internal observing eye that allowed Freud to identify this system in the first place, and we have an octonionic, tetrahedral mirrorhouse.

Finally, let us consider the subjective experience of meditation, as discussed e.g. in (Austin, 1999). Here we have "consciousness without an object" (Merrell-Wolf, 1983), which may be understood as the infusion of the mental mirrorhouse with attention but not content. Each mirror is reflecting the others, without any image to reflect except the mirrors themselves.

Mirroring in Development

Another naturally arising question regards the origin of the mental mirrorhouse faculty, both evolutionarily and developmentally. In both cases, the obvious hypothesis is that during the course of growth, the inner mirrorhouse gains the capability for using more and more mirrors. First comes the capability to internally mirror an external agent; then comes the capability to internally encapsulate an inter-observation process; then comes the capability to internally observe an inter-observation process; then comes the capability to internally observe the observation of an inter-observation process. Of course, the hierarchy need not terminate with the octonionic mirrorhouse; but qualitatively, our suggestion is that levels beyond the octonionic may generally beyond the scope of what the human brain/mind needs to deal with given its limited environment and computational processing power.

To get a better grip on this posited growth process, let us return to Rizzolatti and Arbib's hypothesized role for mirror neurons in language learning. Their stage S5, as described above, involves "proto-signs," which may have initially consisted of pantomime used indirectly (i.e. used, not necessarily to denote specific motor actions, but to denote other activities loosely resembling those motor actions). A mental mirrorhouse corresponding to proto-signs may be understood to comprise 3 mirrors

- Observer

- Pantomimer (carrying out manual actions)
- Object of pantomime (carrying out non-manual actions)

The hypothesis becomes that, via recollecting instances of pantomime using a quaternionic mirrorhouse, the mind imprints pantomimes on the long-term memory, so that they become part of the unconscious in a manner suitable to encourage the formation of new and richer pantomimes.

In general, going beyond the particular example of pantomime, we may posit a quaternionic mirrorhouse corresponding to

- Observer
- Symbols
- Referent

The addition of a fourth mirror then corresponds to ***reflection on the process of symbolization***, which is not necessary for use of language but is necessary for conscious creation of language, as is involved for instance in formalizing grammar or creating abstract mathematics.

There is a clear and fascinating connection here with Piagetan developmental psychology (Piaget and Inhelder, 2000), in which the capability for symbolization is posited to come along with the "concrete operational" stage of development (between ages 7-14 in the average child); and the capability for abstract formal reasoning comes later in the "formal" stage of development. The natural hypothesis in this connection is that the child's mind during the concrete operational stage possesses only a quaternionic mirrorhouse (or at least, that only the quaternionic mirrorhouse is highly functional at this stage); and that the advent of the formal stage corresponds to the advent of the octonionic mirrorhouse.

This hypothesis has interesting biological applications, in the context of the previously hypothesized relationship between mirror neurons and mental mirroring. In this case, if the hypothesized correspondence between number-of-mirrors and developmental stages exists, then it should eventually be neurologically observable via studying the patterns of interaction of cell assemblies whose dynamics are dominated by mirror neurons, in the brains of children at different stages of cognitive development. As noted above, however, experimental neuroscience is currently nowhere near being able to validate or refute such hypotheses, so we must wait at least a couple decades before pursuing this sort of empirical investigation.

Conclusion

The path traced in this paper has been a somewhat complex one, but the moral of the story is simple: There are elegant abstract-algebraic symmetries lurking within the social substructures of the self. The internal structure of the self may well be that of a tetrahedral mirrorhouse and related more complex packing structures; and the Fulleresque vision of an iterating dynamical system of adjacent tetrahedral mirrorhouses may well be an accurate model of critical aspects of the emergent cognitive dynamics of societies of social minds.

The fact that these same algebraic/geometric symmetries pervade modern physics is a fascinating one, which may have to do with the interdependence of physical and psychological reality, or may simply reflect the fact that both domains rely on the same basic mathematical symmetries. Further investigation in this direction, both philosophical and scientific, is merited. One possible guide for such investigation is the observation, made in (Goertzel, 2007), that the quaternion and octonion algebras may be derived as a consequence of a more fundamental algebra of "multiboundary forms," combined with an elementary notion of temporality. It seems that quaternions and octonions emerge almost immediately from basic notions of grouping and time, and in this sense may be viewed as metaphysically fundamental, in a way that cuts deeper than any particular applications in physics, psychology, or other disciplines. How this observation impinges on the particularities of mirrorhouse models of mental activity however, remains to be studied in detail.

Finally, it is worthwhile positioning these observations in terms of the philosophy of mind. In *The Hidden Pattern* (Goertzel, 2006), I have proposed a patternist philosophy of mind, in which minds are construed as sets of patterns associated with physical systems. Some of these patterns may be emergent, in the sense that they arise from the combination of a large number of physical components and sub-patterns, and are not at all evident when looking at these components and sub-patterns in isolation. Clearly the abstract-algebraic patterns noted here fall into the emergent category. If the hypothesis given here is correct, the quaternions and octonions form a significant aspect of the emergent mind of human beings and any other social intelligences, yet clearly this is not because these algebras are wired into the neurons or the neural connectivity patterns of the human brain in any direct way. Rather, they emerge naturally from the logic of interactions of inter-observers, regardless of the details of how these inter-observers arise. Mind is largely emergent; and the ideas outlined here exemplify the subtlety and the elegance of some of the patterns that may emerge within it.

References

- Arbib, M. A., Bonaiuto, J., and Rosta, E. (2006) The mirror system hypothesis: From a macaque-like mirror system to imitation. In Proceedings of the 6th International Conference on the Evolution of Language, pages 3--10.
- Austin, James (1999). Zen and the Brain. MIT Press.
- Baez, John (2002). The Octonions, Bull. Amer. Math. Soc. 39, 145-205. Online HTML version at http://math.ucr.edu/home/baez/octonions/
- Barwise, John and John Etchemendy (1989). The Liar. Oxford Press.
- Baudrillard, Jean (1983). Simulations. Foreign Agents Press, New York
- Blakeslee, Sandra (2006). Cells that Read Minds, New York Times, Jan. 10 2006

- Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., & Freund, H.-J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. European Journal of Neuroscience, 13, 400-404.
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C.A., and Rizzolatti, G. (2004) Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. J Cogn. Neurosci. 16: 114-126.
- Carruthers, Peter and Peter K. Smith (1996). Theories of Theories of Mind, Cambridge University Press
- Collins, T.F.T., E. O. Mann, M.R.H. Hill, E. J. Dommett and S. A. Greenfield (2007). Dynamics of neuronal assemblies are modulated by anaesthetics but not analgesics
- Conway, John and Derek Smith (2003). On Octonions and Quaternions, A K Peters, Natick, MA. ISBN 1-56881-134-9
- Dixon, G. (1994). Division Algebras: Octonions, Quaternions, Complex Numbers and the Algebraic Design of Physics
- Finkelstein, D. and E. Rodriguez (1984). Relativity of topology and dynamics, International Journal of Theoretical Physics, 23, 1065-1098
- Freeman, Anthony, Keith Sutherland and Benjamin Libet (2001). The Volitional Brain. Imprint Academic.
- Freud, Sigmund (1962). The Ego and the Id. W.W. Norton
- Fuller, Buckminster (1982). Synergetics. Macmillan Publishing.
- Goertzel, Ben (1994). Chaotic Logic. Plenum Press.
- Goertzel, Ben (1997). From Complexity to Creativity. Plenum Press.
- Goertzel, Ben (2006). The Hidden Pattern. BrownWalker Press
- Goertzel, Ben (2007). Emergence, Time and Clifford Algebras. Electronic Journal of Theoretical Physics. Special Issue on the Physics of Emergence and Organization.
- Hanson, Andrew (2006). Visualizing Quaternions. Morgan Kaufmann.
- Husserl, E. 1970. Logical Investigations 2 vols. London: Routledge and Kegan Paul.
- Kurzweil, Ray (2005). The Singularity Is Near. Viking
- Lohmar, Dieter (2006). Mirror neurons and the phenomenology of intersubjectivity. Phenomenology and the Cognitive Sciences (2006) 5: 5–16
- Merrell-Wolf, Franklin (1983). The Philosophy of Consciousness Without an Object. Three Rivers Press.
- Metzinger, Thomas (2004). Being No One. MIT Press.
- Piaget, John and Barbel Inhelder (2000). The Psychology of the Child. Basic Books.
- Oberman, Lindsay, Edward M. Hubbard, Joseph P. McCleery, Eric L. Altschulera, Vilayanur S. Ramachandran and Jaime A. Pinedad (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. Cognitive Brain Research, Volume 24, Issue 2, July 2005, Pages 190-198
- Ramachandran, V.S. (2006). Mirror Neurons and imitation learning as the driving force behind "the great leap forward" in human evolution. Edge

Foundation,
http://www.edge.org/3rd_culture/ramachandran/ramachandran_p1.html

- Rizzolatti, Giacamo and Laila Craighero (2004). Annu. Rev. Neurosci. 2004. 27:169–92
- Rizzolatti, G. & Arbib, M. A. (1998). Language within our grasp. Trends in Neurosciences, 21, 188-194.
- Rowan, John (1990). Subpersonalities: The People Inside Us. Routledge Press.
- Thompson, Evan (2001). Between Ourselves: Second-Person Issues in the Study of Consciousness. Imprint Academic.