(New Title, 1 Mar 2013)

Hidden Depths of Triangle Qualia

(Previous title, now sub-title)

Theorems About Triangles, and Implications for Biological Evolution and AI

The Median Stretch, Side Stretch, Triangle Sum, and Triangle Area Theorems

Old and new proofs.

Last updated: Please report bugs (A.Sloman@cs.bham.ac.uk)
9 Sep 2012; 24 Sep 2012; 6 Oct 2012; 10 Nov 2012; 18 Nov 2012; 4 Dec 2012; 24 Dec 2012;
3 Jan 2013; 13 Feb 2013; 24 Feb 2013; 26 Feb 2013; 28 Feb 2013;
(File Completely Reorganised March 2013)
1 Mar 2013; 4 Mar 2013; 19 Mar 2013; 5 May 2013; 7 May 2013 (Reorganised)

Installed: 9 Sep 2012
Installed and maintained by Aaron Sloman

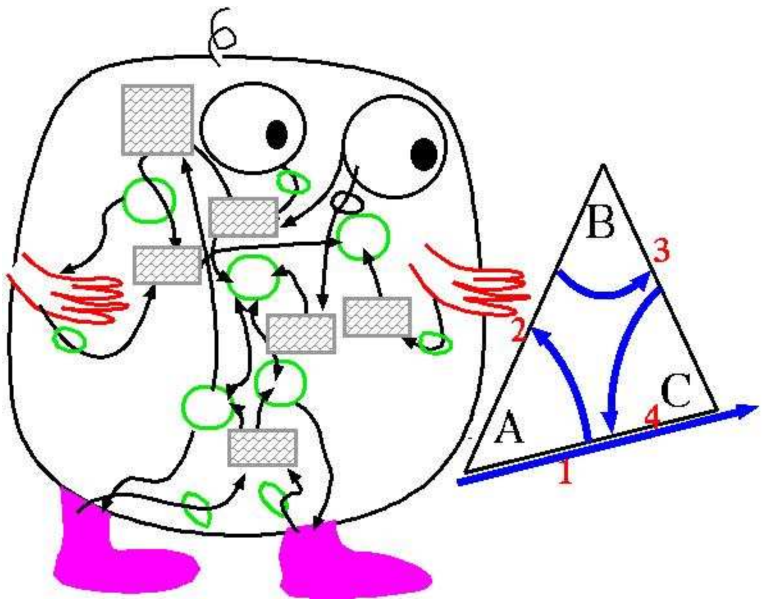
JUMP TO LIST OF CONTENTS

Related documents
This file is http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html
also available as http://tinyurl.com/CogMisc/triangle-theorem.html
A messy PDF version will be automatically generated from time to time:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.pdf
A partial index of discussion notes in this directory is in
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/AREADME.html

See also this discussion of "Toddler Theorems":
http://tinyurl.com/CogMisc/toddler-theorems.html

This document illustrates some points made in a draft, incomplete, discussion of transitions in information-processing, in biological evolution, development, learning, etc. here.
That document and this one are both parts of the Meta-Morphogenesis project, partly inspired by Turing's 1952 paper on morphogenesis.

I suggest below that James Gibson's theory of perception of affordances, is very closely related to mathematical perception of structure, possibilities for change, and constraints on changes (structural invariants). Gibson's ideas are summarised, criticised and extended here:
http://tinyurl.com/BhamCog/talks/#gibson



When will the first baby robot grow up to be a mathematician?

CONTENTS

A Preface was added: 7 May 2013, removed 12 May 2013, now a separate document: Biology, Mathematics, Philosophy, and Evolution of Information Processing http://www.cs.bham.ac.uk/research/projects/cogaff/misc/bio-math-phil.html

---

BACK TO CONTENTS

Introduction: ways to change, or not change, features of a triangle

Aspects of mathematical consciousness of space

This document may appear to some to be a mathematics tutorial, introducing ways of doing Euclidean geometry. It may have that function, but my main aim is to draw attention to products of biological evolution that must have existed before Euclidean geometry was developed and organised in Euclid's Elements over two thousand years ago. I'll give a few examples of apparently very simple human spatial reasoning capabilities concerned with perception of triangles that I think are deeply connected with the abilities of human toddlers and other animals to perceive what James Gibson called "affordances", though I don't think he ever understood the full generality, and depth, of those animal competences.

A core aspect is perceiving what is **possible --** i.e. acquiring information about structures and processes that do not exist but could have existed, and might exist in future -- and grasping some of the **constraints** on those possibilities.

```
(This is not to be confused with discovering probabilities: possibilities
are obviously more basic. The differences between learning about constraints
on what's possible and learning probabilities seem to have been ignored by
most researchers studying probabilistic learning mechanisms, e.g. Bayesian
mechanisms.)
```

The examples I'll present look very simple but have hidden depths, as a result of which there is, as far as I know, nothing in AI that is even close to modelling those animal competences, and nothing in neuroscience that I know of that addresses the problem of explaining how such competences could be implemented in brains. (I am not claiming that computer-based machines **cannot** model them, as Roger Penrose does, only that the current ways of thinking in AI, Computer science, Neuroscience, Cognitive Science, Philosophy of mind and Philosophy of mathematics, need to be extended. I'll be happy to be informed of working models, or even outline designs, implementing such extensions.)

For reasons that will become clearer below this could be dubbed the problem of accounting for **"mathematical qualia"**, or "contents of mathematical/geometric consciousness" -- their evolution, their cognitive functions, and the mechanisms that implement them.

```
I have some ideas about the layers of meta-cognitive, and meta-meta-cognitive
mechanisms that are involved in these processes, which I think are related to
Annette Karmiloff-Smith's ideas about "Representational Redescription", (1992)
but I shall not expand on those ideas here: the purpose of this document is
to present the problem.

For more on this see the Meta-Morphogenesis project:
  http://tinyurl.com/CogMisc/meta-morphogenesis.html
```

An online English version of Euclid's Elements is here:
http://aleph0.clarku.edu/~djoyce/java/elements/elements.html

That was, arguably, the most important, and most influential, book ever written, ignoring highly influential books with mythical or false contents. Unfortunately, this seems to have dropped out of modern education with very sad results.

An Evolutionary Conjecture

My aim here is to provide examples supporting the following conjecture:

> **The discoveries organised and presented in Euclid's Elements were made using products of biological evolution that humans share with several other species of animals that can perceive, understand, reason about, construct, and make use of, structures and processes in the environment-- competences that are also present in pre-verbal humans, e.g. toddlers.**
>
> **Human toddlers, and some other animals, seem to be able to make such discoveries, but they lack the meta-cognitive competences that enable older humans to inspect and reason about those competences, and the discoveries they give rise to.**

I suspect that important subsets of those competences evolved independently in several evolutionary lineages -- including some nest-building birds, elephants, and primates -- because they all inhabit a 3-D environment in which they are able to perceive, understand, produce, maintain, or prevent various kinds of spatial structures and processes. Some of those competences are also present in very young, even pre-verbal, children. But the competences have largely been ignored, or misunderstood, by researchers in developmental psychology, animal cognition, philosophy of mathematics, and more recently AI and robotics. Thinkers who have noticed the gaps sometimes argue that computer-based systems will always have such gaps (e.g. Roger Penrose). That is not my aim, though there is an open question.

```
Research on "tool-use" in young children and other animals usually has
misguided motivations and should be replaced by research on "matter-manipulation"
including use of matter to manipulate matter. But that's a topic for another
occasion.
```

Very often these spatial reasoning competences are confused with very different competences, such as abilities to learn empirical generalisations from experience, and to reason probabilistically. In contrast, this discussion is concerned with abilities to discover what is possible, and constraints on possibilities, i.e. necessities. (These abilities were also noticed by Immanuel Kant, who, I suspect, would have been actively attempting to use Artificial Intelligence modelling techniques to do philosophy, had he been alive now.)

In young humans the mathematical competences discussed here normally become evident in the context of formal education, and as a result it is sometimes suggested, mistakenly, that social processes not only play a role in communicating the competences, or the results of using them, but also determine which forms of reasoning are valid -- a muddle I'll ignore here, apart from commenting that early forms of these competences seem to be evident in pre-school children and other animals, though experimental tests are often inconclusive: we need a deep theory more than we need empirical data.

The capabilities illustrated here are, to the best of my knowledge, not yet replicated in any AI system, though some machines (e.g. some graphics engines used in computer games), may appear to have superficially similar capabilities if their limitations (discussed below) are not exposed. I am not claiming that computers **cannot** do these things, merely that novel forms of representation and reasoning, and possibly new information-processing architectures, will be required -- developing a claim I first made in Sloman(1971), though I did not then expect it would take so long to replicate these animal capabilities. That is partly because I did not then understand the full implications of the claims, especially the connection with some of J.J. Gibson's ideas about the functions of perception in animals discussed below, and the distinction between **online intelligence** and **offline intelligence** also discussed later, which challenges some claims made recently about "embodied cognition" and "enactivism", claims that I regard as deeply confused, because they focus on only a subset of competences associated with being embodied and inhabiting space and time.

> The ideas presented here overlap somewhat with ideas of Jean Mandler on early conceptual development in children and her use of the notion of an "image schema" representation, though she seems not to have noticed the need to account for competences shared with other animals. Studying humans, and trying to model or replicate their competences, while ignoring other species, and the precocial-altricial spectrum in animal development, can lead to serious misconceptions.
> (I am grateful to Frank Guerin for reminding me of Mandler's work, accessible at http://www.cogsci.ucsd.edu/~jean/)
>
> Another colleague recently drew my attention to this paper:
> http://psych.stanford.edu/~jlm/pdfs/Shepard08CogSciStepToRationality.pdf
> Roger N. Shepard,
> The Step to Rationality: The Efficacy of Thought Experiments in Science, Ethics, and Free Will,
> In Cognitive Science, Vol 32, 2008,
>
> It's one thing to notice the importance of these concepts and modes of reasoning. Finding a good characterisation and developing a good explanatory model are very different, more difficult, tasks.

[Note added: 3 Jan 2013]
This document is also closely related to my 1962 DPhil Thesis attempting to explain and defend Immanuel Kant's claim (1781) that mathematical knowledge includes propositions that are necessarily true (i.e. it's impossible for them to be false) but are not provable using only definitions and logic -- i.e. they are not analytic: they are **synthetic necessary** truths.

The thesis is available online in the form of scanned in PDF files, kindly provided by the university of Oxford library:

> Aaron Sloman, *Knowing and Understanding: Relations between meaning and truth, meaning and necessary truth, meaning and synthetic necessary truth*
> http://www.cs.bham.ac.uk/research/projects/cogaff/62-80.html#1962
> The most directly relevant section is Chapter 7 "Kinds of Necessary Truth".
> It is available in faint but readable format in this file
> Also in the Oxford library here.

[End Note]

Very many people have learnt (memorised) the triangle sum theorem, which states that the interior angles of **any** triangle (in a plane) add up to half a rotation, i.e. 180 degrees, or a straight line, even if they have never seen or understood a proof of theorem.

Many who have been shown a proof cannot remember or reconstruct it. I’ll introduce a wonderful proof due to Mary Pardoe later on. For now, notice that the theorem is not at all obvious if you merely look at an arbitrary triangle, such as Figure T:

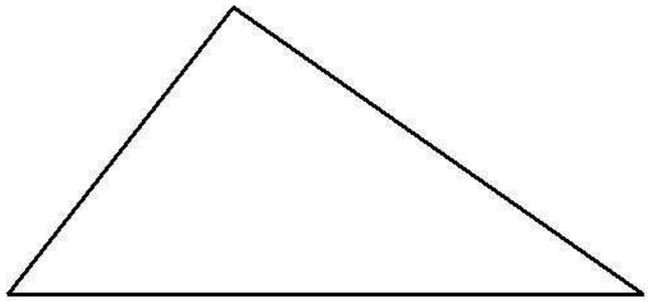


Figure T:

Please stare at that for a while and decide what you can learn about triangles from it. Later you may find that you missed some interesting things that you are capable of noticing.

> I don’t know whether the similarity between this exercise and some of the exercises described by Susan Blackmore in her little book **Zen and the Art of Consciousness,** discussed here, is spurious or reflects some deep connection.

Some ways of thinking about triangles and what can be done with them, including ways of proving the triangle sum theorem, will be presented later. Before that, I’ll introduce some simpler theorems concerning ways of deforming a triangle, and considering whether and how the enclosed area must change when the triangle is deformed.

> **NB:**
> Note that that’s "must change", not "will change", nor "will change with a high probability". These mathematical discoveries are about what **must** be the case. Sometimes researchers who don’t understand this regard mathematical knowledge as a limiting case of empirical knowledge, with a probability of 1.0. Mathematical necessity has nothing to do with probabilities, but everything to do with constraints on possibilities, as I hope will be illustrated below.

Why focus on the human ability to notice and prove some invariant property of triangles? Because it draws attention to abilities to perceive and understand things that are closely related to what James Gibson called "affordances" in the environment, namely: animals can obtain information about possibilities for action and constraints on action that allow actions to be selected and controlled. An example might be detecting that a gap in a wall is too narrow to walk through normally, but not too narrow if you rotate your torso through a right angle and then walk (or sidle) sideways through the gap.

NB:

You can notice the possibility, and think about it, without making use of it. Use of offline intelligence is neither a matter of performing actions at the time or in the immediate future, nor making predictions. The ability to discover such a possibility is not always tied to be able to make use of it in the near future.

Video 6 here illustrates an 19 month old toddler’s grasp of affordances related to a broom, railings and walls: http://tinyurl.com/BhamCog/movies/vid

The video which shows the child manipulating a broom, includes a variety of actions in

which the child seems to understand the constraints on motion of the broom and performs appropriate actions, including moving it so as to escape the restrictions on motion that exist when the broom handle is between upright rails, moving the broom backwards away from a skirting board in order to be able to rotate it so that it can be pushed down the corridor, and changing the orientation of the vertical plane containing the broom so that by the time it reaches the doorway on the right at the end of the corridor the broom is ready to be pushed through the doorway.
NB: I am not claiming that the child understands what he is doing, or proves theorems.

BACK TO CONTENTS

Doing without a global length metric

A feature of such abilities, whose importance should become clearer later, is that they do not depend on the ability to produce or use accurate measurements, using a global scale of length, or area. They do depend on the ability to detect and use ordering information, such as the information that your side-to-side width is greater than the width of a gap you wish to go through, and that your front-to-back width is less than the width of the gap, and your ability to grasp the possibility of rotating and then moving sideways instead of always moving forwards. A **partial** order suffices: you don't need to be able to determine for all gaps viewed at a distance whether your side-to-side or front-to-back distance exceeds the gap.

Using the ordering information (when available), you can infer that although forward motion through the gap is impossible, sideways motion through it is possible. It is important that your understanding is not limited to exactly this spatial configuration (this precise gap width, this precise starting location, this precise colour of shoe, this kind of floor material on which you are standing), since you can abstract away from those details to form a generic understanding of a class of situations in which a problem can arise and can be solved. The key features of the situation are relational, e.g. the gap is narrower than one of your dimensions but greater than another of your dimensions. It does not require absolute measurements. If you learn this abstraction as a child confronted with a particular size gap you can still use what you have learnt as an adult confronted with a larger gap, that the child could have gone through by walking forward.

BACK TO CONTENTS

Offline vs online intelligence

That sort of abstraction to a general schema that can be instantiated in different ways is at the heart of mathematical reasoning (often confused with use of metaphor). It is also important for **offline** intelligence (reasoning and about what can be done, and planning) as opposed to **online** intelligence (used by servo-controlled reactive and homeostatic, systems) a distinction discussed further in another document.

Observation of actions of many different animals, for instance nest building birds, squirrels, hunting mammals, orangutans moving through foliage, and young pre-verbal children, indicates to the educated observer (especially observers with experience of the problems of designing intelligent robots) many animal capabilities apparently based on abilities to perceive, understand and use affordances, some of them more

complex than any discussed by Gibson, including "epistemic affordances" concerned with possible ways of gaining information, illustrated here. For example, If you are in a corridor outside a room with an open door, and you move in a straight line towards the centre of the doorway, you will see more of the room, and will therefore have access to more information, an epistemic affordance.
(This illustrates the connection between theorems in Euclidean geometry and visual affordances that are usable by humans, other animals, and future robots.)

I suspect, but will not argue here, that the human ability to make mathematical discoveries thousands of years ago, that were eventually gathered into a system and published as Euclid’s Elements, depended on the same capability to discover affordances, enhanced by additional meta-cognitive abilities to think about the discoveries, communicate them to others, argue about them, and point out and rectify errors.

(These social processes are important but sometimes misconstrued, e.g. by conventionalist philosophers of mathematics. They will not be discussed here.)

The information-processing (thinking) required in offline intelligence is sometimes too complex to be done entirely within the thinker, and this may have led to the use of external information structures, such as diagrams in sand or clay or other materials, to facilitate thinking and reasoning about the more complex affordances, just as modern mathematicians use blackboards, paper and other external thinking aids, as do engineers, designers, and artists. (As discussed in (Sloman, 1971).)

The roles of external representational media in discovering re-usable generalisations is different from their meta-cognitive role in reasoning about the status of those generalisations, e.g. proving that they are **theorems**. That difference is illustrated but not explained or modelled in this document.

In this document I have chosen some very simple, somewhat artificial, cases, simply to illustrate some of the properties of offline thinking competences, in particular, how they differ from the ability of modern computer simulation engines that can be given an initial configuration from which they compute in great detail, with great precision what will happen thereafter. That simulation ability is very different from the ability to think about a collection of possible trajectories, features they have in common and ways in which they differ, a requirement for the ability to create multi-stage plans. (I am not sure Kenneth Craik understood this difference when he proposed that intelligent animals could use internal models to predict consequences of possible actions, in (Craik, 1943).)

There are attempts to give machines this more general ability to learn about and use affordances, by allowing them to learn and use probability distributions, but I shall try to explain below why that is a very different capability, which lacks the richness and power of the abilities discussed here, though it is sometimes useful. In particular, the probability-based mechanisms lack the ability required to do mathematics and make mathematical discoveries of the kinds illustrated below and in other documents, including "toddler theorems" of the sorts pre-verbal children seem able to discover and use, though there are many individual differences between children: not all can discover the same theorems, nor do they make discoveries in the same order.

One of several motivations for this work is to draw attention to some of what needs to be explained about biological evolution, for the capabilities discussed here all depend on evolved competences -- though some may also be implemented in future machines.

BACK TO CONTENTS

Flaws in enactivism and theories of embodied cognition

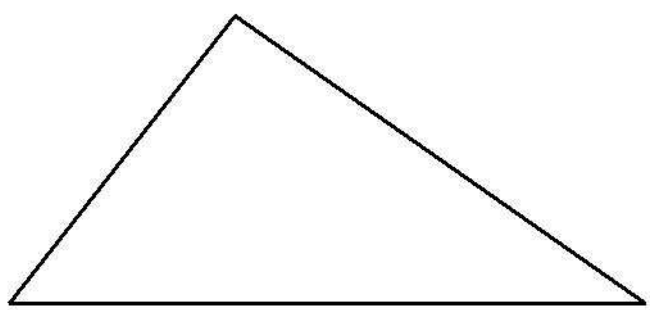
Another motivation is to demonstrate that some researchers in AI/Robotics, cognitive science and philosophy have been seriously misled by the recent emphases on embodiment, and enactivist theories of mind, which are mostly concerned with "online intelligence" and ignore the varieties of "offline intelligence" that become increasingly important as organisms grow larger with more complex and varied needs. Some varieties of offline intelligence (sometimes referred to as deliberative intelligence) required for geometrical reasoning are discussed below. A broader discussion is here.

The small successes of the embodied/enactivist approaches (which are, at best, barely adequate to explain competences of some insects) have diverted attention from the huge and important gaps in our understanding of animal cognition and the implications for understanding human cognition and producing robots with human-like intelligence. Although the online intelligence displayed by the BigDog robot made by Boston Dynamics [REF] is very impressive, it remains insect-like, though perhaps not all insects are restricted to "online" intelligence, which involves reacting to the environment under the control of sensorimotor feedback loops, in contrast with "offline" intelligence, which involves being able to consider, reason about and make use of possibilities (not to be confused with probabilities) some of which are used and some avoided. Related points were made two decades ago by David Kirsh, (though he mistakenly suggests that tying shoelaces is a non-cerebral competence, possibly because it can become one through training, as can many other competences initially based on reasoning about possibilities).

I’ll now return to mathematical reasoning about triangles, hoping that readers will see the connection between that and the ability to use offline intelligence to reason about affordances.

BACK TO CONTENTS

On seeing triangles (again)

Figure T (repeated): 

When you look at a diagram like Figure T, above, you are able to think of it as representing a whole class of triangles, and you can also think about processes that change some aspect of the triangle, such as its shape, size, orientation, or area, in a way that is not restricted to that particular triangle.

How you think about and reason about possible changes in a spatial configuration is a deep question, relevant to understanding human and animal cognition -- e.g. perception of and reasoning about spatial affordances, and also relevant to the task of designing future intelligent machines. Several examples will be presented and discussed below.

As far as I know, there is no current AI or robotic system that can perform these tasks, although many can do something superficially similar, but much less powerful, namely answer questions about, or make predictions about, a very specific process, starting from precisely specified initial conditions. That is not the same as having the ability to reason about an infinite variety of cases. Machines can now do that using equivalent algebraic problems, but they don't understand the equivalence between the algebraic and the geometric problems, discovered by Descartes.

The perception of possible changes in the environment, and constraints on such changes, is an important biological competence, identified by James Gibson as perception of "**affordances**". However, I think he noticed and understood only a small subset of types of affordance. His ideas are presented and generalised in a presentation on his ideas (and Marr's ideas) mentioned above.

I shall present several examples of your ability to perceive and reason about possibilities for change, and constraints on those possibilities inherent in a spatial configuration, extending the discussion in my 1996 "Actual Possibilities" paper.

In particular, we need to discuss your ability to:

a) perceive a shape,
b) notice the possibility of a certain constrained transformation of that shape
c) discover and prove a consequence of that constraint.

Such mathematical competences seem to be closely related to much more wide-spread animal competences involving perception of possibilities for change, including possibilities for action in the environment; and reasoning about consequences of realising those possibilities. The mathematical competences build on these older, more primitive, competences, which seem largely to have gone unnoticed by researchers in human and robot cognition. I have tried to draw attention to examples that can be observed in young children in a discussion of "toddler theorems".

Several proofs of simple theorems will be presented below, making use of your ability to perceive and reason about possible changes in spatial configurations. I'll start with some deceptively simple examples relating to the area enclosed by a triangle.

A developmental neuroscience researcher whose work seems to be closely related to this is Annette Karmiloff-Smith, whose ideas about "Representational Redescription" in her 1992 book "Beyond Modularity" are discussed here.

BACK TO CONTENTS

The "Median Stretch Theorem"

The first theorem concerns the consequences of moving one vertex of a triangle along a median, while the other two vertices do not move. I shall start by assuming that the concept of the **area** enclosed by a set of lines is understood, and that at least in some cases we can tell which of two areas is larger. Later, I'll return to hidden complexity in the concept of area.

A median of a triangle is a straight line between the midpoint of one side of the triangle to the opposite vertex (corner). The dashed arrows in Figure M (a) and (b) lie on medians of the triangles composed of solid lines. The dashed arrows in triangles (a) and (b) have both been extended beyond the median, which terminates at the vertex. The dotted lines indicate the new locations that would be produced for the sides of the triangle if the vertex were moved out, as shown.

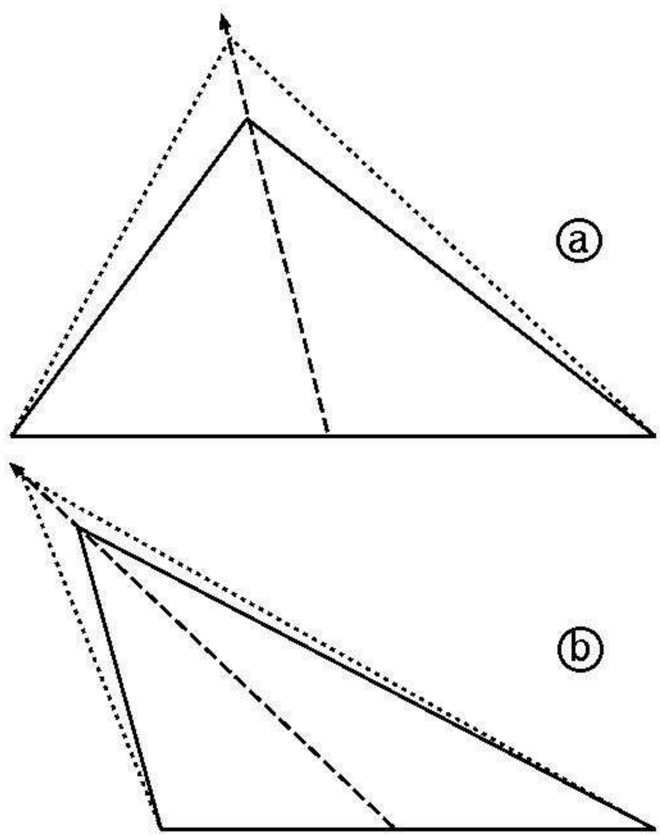


Figure M:

Consider what happens if we draw a median in a triangle, namely a line from the midpoint of one side through the opposite vertex, and then move that vertex along the extension of the median, as shown in figure M(a). You should find it very obvious that moving the vertex in one direction along the median increases the area of the triangle, and movement in the other direction decreases the area. Why?

We can formulate the "Median stretch theorem" (MST) in two parts:

(MST-out)
IF a vertex of a triangle is moved along a median **away from** the opposite side,
THEN the area of the triangle **increases**.

(MST-in)
IF a vertex of a triangle is moved along a median **towards** the opposite side,
THEN the area of the triangle **decreases**.

As figure M(b) shows, it makes no difference if the vertex is not perpendicularly above the opposite side: the diagrammatic proof displays an invariant that is not sensitive to alteration of the initial shape of the triangle, e.g. changing the slant of the median, and changing the initial position of the vertex in relation to the opposite side makes no difference to the truth of theorem. Why?

A problem to think about:
How can you be sure that there is no counter-example to the theorem, e.g. that stretching or rotating the triangle, or making it a different colour, or painting it on a different material, or transporting it to Mars, will not make any difference to the truth of (MST-out) or (MST-in)?

NOTE:
As far as I know the median stretch theorem has never been stated previously, though I suspect it has been used many times as an "obvious" truth in many contexts, both mathematical and non-mathematical.
If you know of any statement or discussion of the theorem, please let me know.

Note added 24 Feb 2013:
Readers may find it obvious that the median stretch theorem is a special case of a more general stretch theorem that can be formulated by relaxing one of the constraints on the lines in the diagram. Figuring out the generalisation is left as an exercise for the reader. (Feel free to email me about this.) Compare (Lakatos, 1976).

Added 13 Feb 2013
Julian Bradfield pointed out, in conversation, that one way to think about the truth of MST-OUT is to notice that the change of vertex **adds two triangles** to the original triangle. Likewise, in support of MST-IN, moving the vertex inwards **subtracts** two triangles from the original area.
(Below I suggest decomposing the proof into two applications of the Side-Stretch-Theorem (SST), which can also be thought of as involving the addition or subtraction of a triangle.)

BACK TO CONTENTS

NOTE: How can areas be compared?

The concept of "area" used here may seem intuitive and obvious, but generalising it to figures with arbitrary boundaries is far from obvious and requires the use of sophisticated mathematical reasoning about limits of infinite sequences.

For example, how can you compare the areas of an ellipse and a circle, neither of which completely encloses the other? What are we asking when we ask whether the blue circle or the red ellipse has a larger area in Figure A, below?

It is obvious that the black square contains less space than the blue circle, and also contains less space than the red ellipse, simply because all the space in the square is also in side the circle and inside the ellipse. But what does it mean to ask whether one object contains more space than another if each cannot fit inside the other?

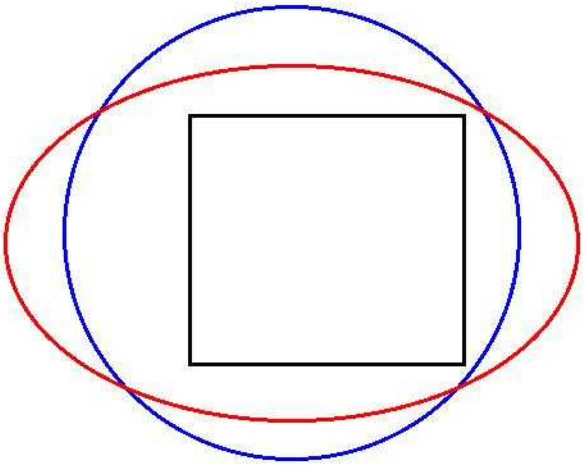


Figure A:

Some teachers try to get young children to think about this sort of question by cutting out figures and weighing them. But that assumes that the concept of weight is understood. In any case we are not asking whether the portion of paper (or screen!) included in the circle weighs more than the portion included in the ellipse. There is a correlation between area and weight (why?) but it is not a reliable correlation.
Why not?

The standard mathematical way of defining the area of a region includes imagining ways of dividing up non-rectangular regions into combinations of regions bounded by straight lines (e.g. thin triangles, or small squares), using the sum of many small areas as an approximation to the large area. The smaller the squares the better the approximation, in normal cases.
(Why? -- Another area theorem).

For our purposes in considering the triangles in Figure M and Figure S, most of those difficulties can be ignored, since we can, for now, use just the trivial fact that if one region totally encloses another then it has a larger area than the region it encloses, leaving open the question of how to define "area", or what it means to say that area A1 is larger than area A2, when neither encloses the other. Our theorems about stretching (MST above and SST below) only require consideration of area comparisons when one area completely encloses another.

Cautionary note:
It is very easy for experimental researchers studying animals or young children to ask whether they do or do not understand areas (or volumes, or lengths of curved lines), and devise tests to check for understanding, without the researchers themselves having anything like a full understanding of these concepts that troubled many great mathematicians for centuries. (I have checked this by talking to some of the researchers, who had not realised that the

resources for thinking about areas and volumes in very young children might support only a **partial** ordering of areas.

These problems are usually made explicit only to students doing a degree in mathematics.

BACK TO CONTENTS

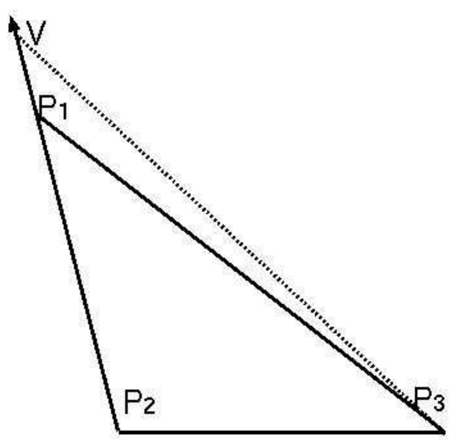
Using the Side Stretch Theorem to prove the Median Stretch Theorem

In this section, we'll introduce the Side Stretch Theorem (SST) and show how it was implicitly assumed in the proof of the Median Stretch Theorem, above.

If you think about why the MST must always be true, you may notice that the problem can be broken down into two parts, because the original triangle has two triangular parts, one on each side of the median; and moving the vertex along the median always either increases the area of each part or decreases the area of each part. If the area of each **part** of the triangle is increased by a movement of the vertex, then the area of the **whole** triangle must be increased. Likewise for a decrease in the area of each part.
(See discussion above on "What does 'area' mean?")

If you think about your reasoning about the change in area of each of the two sub-triangles, you may notice another theorem, which could be called "The side stretch theorem" illustrated in figure S.

Figure S:



We can formulate the "Side stretch theorem" (SST) in two parts:

(SST-out)
IF a vertex of a triangle is moved along an extended side **away** from the interior of the side (as in Figure S)
THEN the area of the triangle increases.

(SST-in)
IF a vertex of a triangle is moved along a side **towards** the interior of that side,
THEN the area of the triangle decreases.
(Draw your own figure for this case.)

Comparing Figure S, with Figure M (a) or Figure M (b) should make it clear that when a vertex moves along the median of either of the triangles in Figure M, then there are also two smaller triangles, each of which has one side on the median, and when the vertex of the big triangle moves along the median then the (shared) vertex of each of the smaller triangles moves along the shared side.

Moreover, when the shared vertex in Figure M (a) or (b) moves along the median, both of the smaller triangles either increase decrease in area, simultaneously, from which it follows that their combined area must increase when the vertex moves along the median **away from** the opposite side and decrease when the vertex moves along the median **towards** the opposite side.

For now, I'll leave open the question whether the Side Stretch Theorem (SST-in/out) can be derived from something more basic and obvious. Instead, let's consider what we mean by "area" before returning to shape changes and their consequences.

BACK TO CONTENTS

The theorem, and its proof, involves continuity and discontinuity

As the vertex moves **further** from the other end of the side in question, the area will continuously increase. However, if the vertex moves in the opposite direction, **towards** the other end, then the change in direction of motion necessarily induces a change in what happens to the area: instead of **increasing**, the area must **decrease**. Although the vertex can move up and down with continuously changing velocity, or even continuously changing acceleration, there are unavoidable **discontinuities**: the direction of motion can change, and so can whether the area is increasing or decreasing.

It is also the case that during continuous motion in the same direction a "virtual discontinuity" can occur. If the vertex starts beyond the original position and moves back towards the other end of the line, then the area will be continuously decreasing. But for an observer that has stored information about the original position, or for that matter any other position on the line, there will be a discontinuous change from an area greater than the original area to an area less than the original area -- with "instantaneous equality" separating the two phases of motion. This discontinuity is not intrinsic to the motion, but involves an external relationship to a previous state. There are many cases where understanding mathematical relationships or understanding affordances involves being able to detect such relational discontinuities (phase changes of a sort).

The relationship between direction of motion of the vertex and whether the area increases or decreases can be seen to be an **invariant** relationship. But it is not clear what information-processing mechanisms make it possible to discover that invariance, or necessity. Notice that this is utterly different from the kind of discovery currently made by collecting large numbers of observations and then seeking statistical relationships in the data generated, which is how much robot learning is now done. The kind of learning described here, when done by a human does not require large amounts of data, nor use of statistics. There are no probabilities involved, only invariant relationships: if the perpendicular distance increases the area must also.

BACK TO CONTENTS

Another way of modifying a triangle

Instead of considering what happens when we move the upper vertex in Figure T so that it moves along a median, we can consider possible changes in which the vertex remains at the **same** distance from the opposite side, which would be achieved by moving it in a line **parallel** to the opposite side instead of a line **perpendicular** to the opposite side. (The notion of parallelism includes subtleties that will be ignored for now)

In Figure Para, below, two new dotted triangles have been added, a red one and a blue one, both with vertices on the dashed line, parallel to the base of the original triangle, and both sharing a side (the base) with the original triangle.

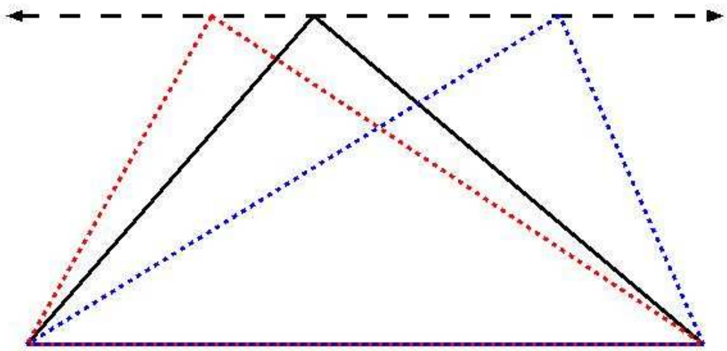


Figure Para:

The figure shows that moving the top vertex of the original triangle along a line **parallel** to the opposite side will definitely not produce a triangle that encloses the original, because, whichever way the vertex is moved on that line (the dashed line in Figure Para) the change produces a triangle with two new sides, one partly inside the old triangle and the other outside the old triangle. So the new triangle cannot enclose the old one, or be enclosed by it.

Proving the theorem that moving a vertex of a triangle in a direction **parallel** to the opposite side does not alter the area is left as an exercise for the reader, though I shall return to it below.

There is a standard proof used to establish a formula for the area of a triangle, which requires consideration of different configurations, as we'll see below. (The need for case analysis is a common feature of mathematical proof Lakatos 1976).

Exercise for the reader:

```
Try to formulate a theorem about what happens to the sides of a triangle if a vertex
moves along a line that goes through the vertex but does not go through the triangle,
like the dashed line in Figure Para, above.
```

BACK TO CONTENTS

Some observations on the above examples

The examples above show that many humans looking at a triangle are not only able to see and think about the particular triangle displayed, but can also use the perceived triangle to support thinking and reasoning about large, indeed infinitely large, sets of **possible** triangles, related in different ways to the original triangle.

Note:

The concept of an infinitely large set being used here is subtle and complex and (as Immanuel Kant noted) raises deep questions about how it is possible to grasp such a concept. For the purposes of this discussion it will suffice to note that if we are considering a range of cases and have a means of producing a new case different from previously considered cases, then that supports an unbounded collection of cases.

For example, in Figure S, where a vertex of a triangle is moved along an extension of a side of the triangle, between any two positions of the vertex there is at least one additional possible position, and however far along the extended side the vertex has been moved there are always further locations to which it could be moved.

So anyone who is squeamish about referring to infinite sets can, for our purposes, refer to unbounded sets.

Below I'll discuss some implications for meta-cognition in biological information processing.

In some cases the new configurations thought about include additional geometrical features, specifying constraints on the new triangle, for example the constraint that a vertex is on a median, or extension of a median, of the original triangle, or on a particular line parallel to one of the sides. Such constraints, involving lines or circles or other shapes can be used to limit the possible variants of the original shape, while still leaving infinitely many different cases to be considered.

However, the infinity of possibilities is reduced to a small number of cases by making use of common features, or invariants, among the infinity of cases.

For instance the common feature may be a vertex lying on a particular line, such as a median of the original triangle (as in the Median Stretch Theorem (MST) above, or an extended side of the original triangle (as in the Side Stretch Theorem (SST above)). Then we can divide the infinity of cases of change of length to two subsets: a change that increases the length and a change that decreases the length, as was done for each of the theorems. Each subset has an invariant that can be inspected by a perceiver or thinker with suitable meta-cognitive capabilities, discussed further below.

There are more complex cases, as we'll see when considering vertices lying along a perpendicular to one of the sides of the triangle, required for proving a theorem about how to calculate the area of a triangle. The complication is that there are infinitely many perpendiculars to a given line, whereas there is only one extension to the line, as shown in Figure M and Figure S. However, in all figures required for discoveries of the sorts we are discussion, there is an additional infinity of cases because of possible variations in the original triangle considered, before effects of motion of the vertex are studied.

BACK TO CONTENTS

The role of meta-cognition

This ability to think about infinitely many cases in a finite way seems to depend on the biological **meta-cognitive ability** to notice that members of a set of perceived structures or processes share a common feature that can be described in a meta-language for describing spatial (or more generally perceptual) information structures and processes. An example would be noticing that between **any** two stages in a process there are intermediate stages, and that between any two locations on a line, thicknesses of a line, angles between lines, amounts of curvature, there are **always** intermediate cases, with the implication that there are intermediate cases between the intermediate cases and the intermediate cases never run out.

> **NOTE**: for now we can ignore the difference between a set being dense and being continuous -- a difference that mathematicians did not fully understand until the 19th Century. I shall go on referring loosely to ’continuity’ to cover both cases.

This ability to notice that some perceived structure or process is continuous, and therefore infinite, is meta-cognitive insofar as it requires the process of perceiving, or imagining, a structure or process to be monitored by another process which inspects the changing information content of what is being perceived, or imagined, and detects some feature of this process such as continuity, or such as being divisible into discrete cases (e.g. motion away from or towards a line). A more complex meta-cognitive process may notice an invariant of the perceived structure or process, for instance detecting that a particular change necessarily produces another change, such as increasing area, or that it preserves some feature, e.g. preserving area.

> **NOTE:** The transitions in biological information processing required for organisms to have this sort of meta-cognitive competence have largely gone unnoticed. But I suspect they form a very important feature of animal intelligence that later provided part of the basis for further transitions, including development of meta-meta-meta... competences required for human intelligence. (Chappell&Sloman 2007)

These meta-cognitive abilities are superficially related to, but very different from, abilities using statistical pattern recognition techniques to cluster sets of measurements on the basis of co-occurrences. Examples of non-statistical competences include being able to notice that certain differences between cases are irrelevant to some relationship of interest, or being able to notice a way of partitioning a continuous set of cases into two or more non-overlapping sub-sets, possibly with partially indeterminate (fuzzy) boundaries between them. In contrast, many of the statistical techniques require use of large numbers of precise measures in order to detect some pattern in the collection of measures (e.g. an average, or the amount of deviation from the average, or the existence of clusters).

For example, you should find it obvious that the arguments used above based on Figure M and Figure S to prove the Median Stretch and Side Stretch theorems (MST and SST) do not depend on the sizes or shapes of the original triangles. So the argument covers infinitely many different triangular shapes. The features that change if a vertex is moved away from the opposite side along a median or along a side will always change in the same direction, namely, increasing the area.

Noticing an invariant topological or geometrical relationship by abstracting away from details of one particular case is very different from searching for correlations in a large number of particular cases represented in precise detail. For example computation of averages and various other statistics requires availability of many particular, precise, measurements, whereas the discovery process demonstrated above does not require even one precisely measured case. The messy and blurred Figure S-b will do just as well to support the reasoning used in connection with the more precise Figure S, though even that has lines that are not infinitely thin.

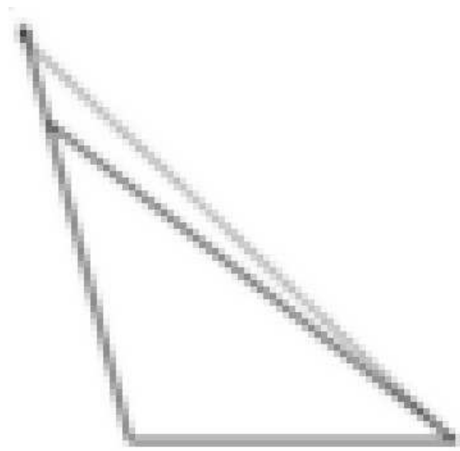


Figure S-b:

Most of these points were made, though less clearly in (Sloman, 1971) which also emphasised the fact that for mathematical reasoning the use of external diagrams is sometimes essential because the complexities of some reasoning are too great for a mental diagram. (These points were generalised in Sloman 1978 Chapter 6). Every mathematician who reasons with the help of a blackboard or sheet of paper knows this, and understands the difference between using something in the environment to **reason** with and using physical apparatus to do **empirical** research, though it took some time for many philosophers of mind to notice that minds are extended. (The point was also made in relation to reference to the past in P.F. Strawson's 1959 book, Individuals, An essay in descriptive metaphysics.)

NOTE ADDED 12 Sep 2012: DIAGRAMS CAN BE SLOPPY

In many cases a mathematician constructing a proof will draw a diagram without bothering to ensure that the lines are perfectly straight, or perfectly circular, etc., or that they are infinitely thin (difficult with line drawing devices available to us). That's because what is being studied is not the particular physical line or lines drawn on paper or sand, etc. The lines drawn are merely representations of perfect Euclidean lines whose properties are actually very different, and very difficult to represent accurately on a blackboard or on paper. E.g. drawing an infinitely thin line has been a problem.

In fact, the lines don't need to be drawn physically at all: they can be imagined and reasoned about, though in some cases a physical drawing can help with both memory and reasoning.

BACK TO CONTENTS

Perception of affordances

All this seems to b closely related to the ability of animals to perceive affordances of various kinds, as discussed in http://tinyurl.com/BhamCog/talks/#gibson .

In particular, the kind of mathematical reasoning about infinite ranges of possibilities and implications of constraints, seems to be closely related to the ability of young children and other animals to discover possibilities for change in their environments, and abilities to reason about invariants in subsets of possibilities that can be relied on when planning actions in the environment.
This leads to the notion of a "toddler theorem" discussed in
http://tinyurl.com/BhamCog/talks/#toddler
http://tinyurl.com/CogMisc/toddler-theorems.html

I suspect that the reasoning using schematic diagrams illustrated above, and also illustrated below in Pardoe's proof of the Triangle Sum Theorem (TST), shares features with animal reasoning about affordances, in which conclusions are reliably drawn about invariants that are preserved in a process, or about impossibilities in some cases -- e.g. it is impossible to completely enclose a bounded area using only two straight lines. Why is it impossible?

There have been attempts to simulate mathematical reasoning using diagrams by giving machines the ability to construct and run simulations of physical processes. But that misses the point: a computer running a simulation in order to derive a conclusion can handle only the specific values (angles, lengths, speeds, for example) that occur in the initial and predicted end states when the simulation runs. Moreover, the simulation mechanisms have to be carefully crafted to be accurate. In contrast, as pointed out above, a human reasoning about a geometrical theorem does not require precision in the diagrams and the conclusion drawn is typically not restricted to the particular lengths, angles, areas, etc. but can be understood to apply to infinitely many different configurations satisfying the initial conditions of the proof.
(See the recent discussion between Mary Leng and Mateja Jamnik, in The Reasoner.)

This seems to require something very different from the ability to run a simulation: it requires the ability to manipulate an **abstract representation** and to interpret the results of the manipulation in the light of the representational function of the representations manipulated. In other words mathematical thinking using diagrams and imagined transformations of geometrical structures, as illustrated above, inherently requires meta-cognitive abilities to notice and reason about features of a process in which semantically interpreted structures are manipulated. The noticing and reasoning need not itself be noticed or reasoned about, although that may develop later (as seems to happen, in different degrees, in humans).

I suspect that many animals, and also pre-verbal human children have simplified versions of that ability, but do not know that they have it. They cannot inspect their reasoning, evaluate it, communicate it to others, wonder whether they have covered all cases, etc. There seems to be a kind of meta-cognitive development that occurs in humans, perhaps partly as a result of learning to communicate and to think using an external language. It may be that some highly intelligent non-human

reasoners have something closely related. But we shall need more detailed specifications of the reasoning processes and the mechanisms required, before we can check that conjecture.
(Annette Karmiloff-Smith's ideas about "Representational Redescription", in "Beyond Modularity" are also relevant.)

We also need more detailed specifications in order to build robots with these "pre-historic" mathematical reasoning capabilities -- which, as far as I can tell, no AI systems have at present. Unlike Roger Penrose, who seems to me to have noticed similar features of mathematical reasoning, I don't think there is any obvious reason why computer based systems cannot have similar capabilities. However it may turn out that there is something about animal abilities to perceive, or imagine, processes of continuous change at the same time as noticing logically expressible constraints or invariants of those processes that requires information processing mechanisms that have so far not been understood. Alternatively, it may simply be that no high calibre AI programmers have attempted to implement competences of the sorts required to invent and understand Pardoe's proof, or many of the traditional proofs used in Euclidean geometry.

These ideas suggest a host of possible investigations of ways in which human capabilities change, along with the reasoning competences of intelligent animals such as squirrels, elephants, apes, cetaceans, octopuses, and others.

BACK TO CONTENTS

The Triangle Sum Theorem

The triangle sum theorem is normally expressed as "The interior angles of a triangle add up to 180 degrees". This assumes a standard way of measuring angles, according to which a complete rotation would be 360 degrees and a half rotation 180 degrees. But we can equivalently express the theorem as "The interior angles of a triangle add up to a straight line, which does not require any conventional unit for measuring angles". As we'll see below that suggests a way of proving the theorem by considering a succession of rotations and seeing what they add up to, an idea suggested by Mary Pardoe when she was a mathematics teacher.

There is a standard way (or small set of standard ways) of proving the theorem

> **Triangle Sum Theorem (TST):** The interior angles of a triangle add up to a straight line, or half a rotation (180 degrees).

These standard methods all make use of some version of Euclid's parallel postulate, (Axiom 5 in Euclid's elements) which can be formulated in several equivalent ways, e.g.

> **Definition:**
> Two straight lines L1 and L2 are parallel if and only if they are co-planar and have no point in common, no matter how far they are extended.

> **Postulate:**
> Given a straight line L in a plane, and a point P in the plane not on L, there is exactly one line through P that is in the plane and parallel to L.

All of this presupposes the concept of "straightness" of a line. For now I'll take that concept for granted, without attempting to define it, though we can note that if a line is straight it is also symmetric about itself (it coincides with its reflection) and also it can be slid along itself without any gaps appearing. If it were possible to view a straight line from one end it would appear as a point.

The "standard" ways of proving the TST make use of properties of angles formed when a straight line joins or crosses a pair of parallel lines:

COR: Corresponding angles are equal:
If two lines L1, L2 are parallel and a third line L3 is drawn from any point P1 on L1 to a point P2 on L2 and continued beyond P2,
then the angle that L1 makes with the line L3 at point P1, and the angle L2 makes with the line L3 at point P2 (where the angles are on the same side of both lines) are equal.

ALT: Alternate angles are equal:
If two lines L1, L2 are parallel and a third line L3 is drawn from any point P1 on L1 to a point P2 on L2,
then the angle L1 makes with the line L3 at point P1, and the angle L2 makes with the line L3 at point P2 (on the opposite sides of both lines) are equal.

For more on transversals and relations between the angles they create, see
http://www.mathsisfun.com/geometry/parallel-lines.html
That page teaches concepts with some interactive illustrations, but presents no proofs.

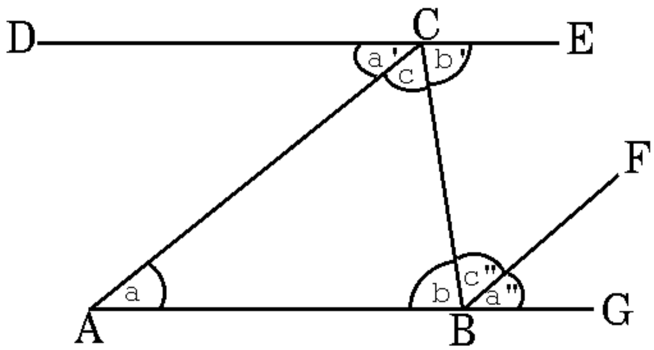
The Euclidean proofs of COR and ALT are presented here:
http://www.proofwiki.org/wiki/Parallel_Implies_Equal_Alternate_Interior_Angles,_Corresponding_Angles,_and_Supplementary_Interior_Angles

BACK TO CONTENTS

The "standard" proofs of the "Triangle Sum Theorem"

Two "standard" proofs of the triangle sum theorem using parallel lines, and the Euclidean theorems **COR** and/or **ALT** stated above, are shown below in Figure Ang1:

To be proved:
In triangle ABC the interior angles, a, b, c sum to a straight line.



Proof 1: (Does not use line BF or angles a" c")
Line DE is parallel to line AB, and AC is a transversal joining them. Since a and a' are alternate angles, they must be equal; likewise BC is a transversal joining parallel lines AB and DE, with b and b' alternate angles and therefore equal.

So: a+b+c = a'+b'+c Q.E.D

Proof 2: (Does not use line DE or angles a' b')
Line BF is drawn parallel to AC. Line AB, extended to G is a transversal crossing parallel lines AC and BF. So a and a" are corresponding angles and therefore equal. CB is another transversal joining the two parallel lines AC and BF with c and c" alternate angles and therefore equal.

Figure Ang1: So: a+b+c = a"+b+c" Q.E.D

Warning: I have found some online proofs of theorems in Euclidean geometry with bugs apparently due to carelessness, so it is important to check every such proof found online. The fact that individual thinkers can check such a proof is in part of what needs to be explained.

BACK TO CONTENTS

Mary Pardoe’s proof of the Triangle Sum Theorem

Many years ago at Sussex university I was visited by a former student Mary Pardoe (nee Ensor), who had been teaching mathematics in schools. She told me that her pupils had found the standard proof of the triangle sum theorem hard to take in and remember, but that she had found an alternative proof, which was more memorable, and easier for her pupils to understand.

Her proof just involves rotating a single directed line segment (or arrow, or pencil, or ...) through each of the angles in turn at the corners of the triangle, which must result in its ending up in its initial location pointing in the opposite direction, without ever crossing over itself.

So the total rotation angle is equivalent to a straight line, or half rotation, i.e. 180 degrees, using the convention that a full rotation is 360 degrees.

The proof is illustrated below in Figure Ang2.

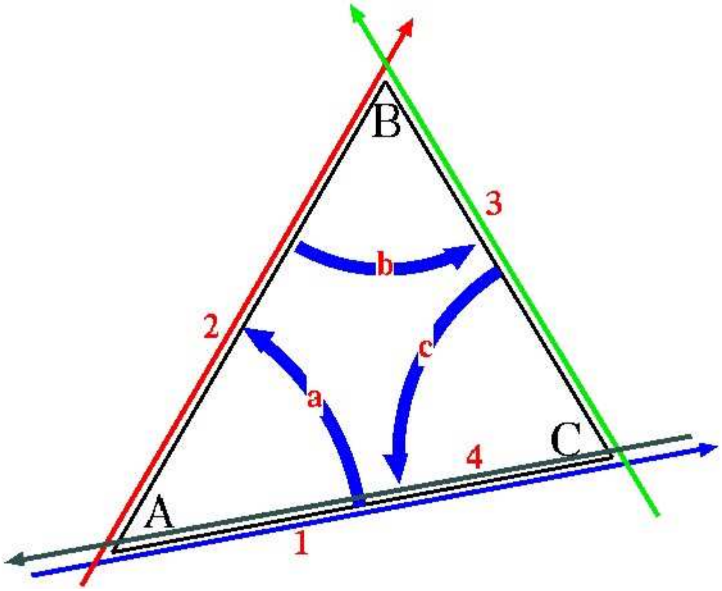


Figure Ang2:

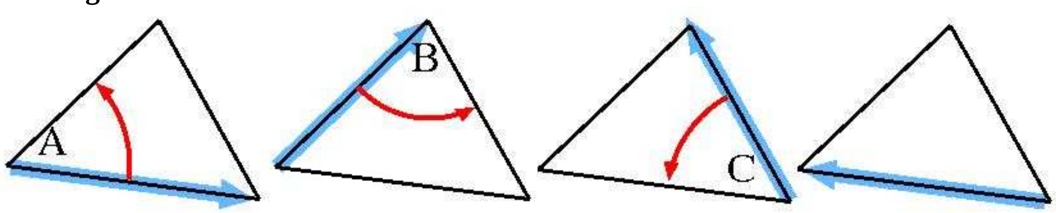
In order to understand the proof, think of the blue arrow, labelled "1", as starting on line AC, pointing from A to C, and then being rotated first around point A, then point B, then point C until it ends up on the original line but pointing in the direction of the dark grey arrow, labelled "4".

So, understanding the proof involves considering what happens if

- the blue arrow labelled "1" initially lies on the side AC of the triangle,
- then is rotated counter-clockwise through angle **A**, indicated by the curved arrow labelled "a", to the location of the red arrow labelled "2", assumed to lie along the side AB of the triangle,
- then rotated counter-clockwise through angle **B**, as indicated by curved arrow "b", to the location of the green arrow labelled "3", assumed to lie along the side BC of the triangle,
- then rotated counter-clockwise through angle **C**, as indicated by curved arrow "c", to the location of the dark grey arrow labelled "4" assumed to lie along the side CA of the triangle.

A "time-lapse" presentation of the proof may be clearer, as shown in Figure Ang3:

Figure Ang3:



It may be best to think of the proof not as a static diagram but as a **process**, with stages represented from left to right in Figure Ang3. In the first stage, the pale blue arrow starts on the bottom side of the triangle, pointing to the right then is rotated through each of the internal angles A, B, C, always rotated in the same direction (counter-clockwise in this case), so that it lies on each of the other sides in succession, until it is finally rotated through the third angle, c, after which it lies on the original side of the triangle, but obviously pointing in the opposite direction. Some people may prefer to rotate something like a pencil rather than imagining a rotation depicted by snapshots.

In this triangle the sides are not very different in length, which conceals a problem that can arise if the first side the arrow is on is very short and the other two are longer. If the length of the arrow is fixed by the length of the first side, you would need to imagine either that the arrow stretches or shrinks as it rotates, or that it slides along a line after reaching it so as to be able to rotate around the next vertex. Alternatively you can imagine that the depicted arrow is part of a much longer invisible arrow, so that, as the invisible arrow rotates from one side to another, it always extends beyond both ends of the new side, and can then rotate around the next vertex. I leave it to the reader to think about these alternatives and what difference they make to the proof, and to the cognitive competences required to construct and understand the proof.

For an arrow to be rotated in a plane and end up lying in its original position it must have been rotated through some number of half-rotations. (Each half rotation brings it back to the original orientation, but pointing in alternate directions.)

Since (1) the arrow at no point crossed over its original orientation, and (2) it ended up pointing in the opposite direction to its original orientation, the total rotation was through a half circle -- which is clear if you actually perform the rotations using a physical object, such as a pencil.

And since that rotation was made up of combined rotations through angles A, B, and C, those three angles must add up to a half circle, i.e. 180 degrees.

A crucial feature of our ability to think about the diagram and the process, is that we (presumably including you, the reader) can see that the key features of the process could have been replicated, no matter what the size or orientation of the triangle, no matter what the lengths of the sides or the sizes of the angles, no matter which side the arrow starts on, no matter which way it is pointing initially, and no matter in which order the rotations are performed, e.g. A then B then C, or C reversed, then B reversed, then A reversed.

This proof of the triangle sum theorem, using a rotating moving arrow, works for all possible triangles on a plane -- as do the standard Euclidean proofs using parallel lines.

This proof is unlike standard proofs in Euclidean geometry since it involves consideration of continuous processes, and therefore involves time and temporal ordering, whereas Euclidean geometry does not explicitly mention time or processes -- though there are some theorems about the locus of point or line satisfying certain constraints, which can be interpreted either as specifying properties of processes extended in time, or as properties of static trajectories, e.g. properties of lines or curves.

NOTE:

> http://tinyurl.com/CogMisc/p-geometry.html presents a more detailed, but still incomplete, discussion, of the geometrical prerequisites for some of the above reasoning. It introduces the idea of P-geometry, which is intended to be Euclidean geometry without the Axiom of Parallels (Euclid's Axiom 5), but with time and motion added, including translation and rotation of rigid line-segments. When I get time, I should add the side-stretch theorem to it (SST above).

BACK TO CONTENTS

Is the Pardoe proof valid?

NOTE: I have presented Mary Pardoe's proof in several places, over several years, e.g.

```
Aaron Sloman, 2008,
Kantian Philosophy of Mathematics and Young Robots,
in Intelligent Computer Mathematics,
Eds. Autexier, S., Campbell, J., Rubio, J., Sorge, V., Suzuki, M., and Wiedijk, F.,
LLNCS no 5144, pp. 558-573, Springer,
http://www.cs.bham.ac.uk/research/projects/cosy/papers#tr0802

Aaron Sloman, 2010,
If Learning Maths Requires a Teacher, Where did the First Teachers Come From?
In Proceedings Symposium on Mathematical Practice and Cognition,
AISB 2010 Convention, De Montfort University, Leicester
http://www.cs.bham.ac.uk/research/projects/cogaff/10.html#1001

And in talks on mathematical cognition and philosophy of mathematics here:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/
```

The presentations produced no responses -- either critical or approving, except that in one informal discussion a mathematician objected that the proof was unacceptable because the surface of a sphere would provide a counter example. However, the surface of a sphere provides no more and no less of a problem for Pardoe's proof than for the standard Euclidean proofs since both proofs are restricted to planar surfaces.

I tried searching for online proofs to see if anyone else had discovered this proof or used it, but nothing turned up. The proof using rotation is so simple and so effective that both Mary Pardoe and I feel sure it must have been discovered previously.

NOTE ADDED 6 Oct 2012:
I have very recently discovered that as a result of the discussion I stirred up in 2010 on the MKM-IG email list, Andrea Asperti mentioned the proof (and the email discussion) in this paper, discussing related issues:

```
Andrea Asperti, Proof, Message and Certificate,
in AISC/MKM/Calculemus, 2012, pp. 17--31,
Online: http://www.cs.unibo.it/~asperti/PAPERS/proofs.pdf
http://dx.doi.org/10.1007/978-3-642-31374-5_2
```

NOTE:
There is a "process" version of the proof of Pythagoras theorem that makes use of a video. A version implemented in Pop-11 is illustrated in the video in this tutorial:
http://tinyurl.com/BhamCog/tutorials/pythagoras.html

The video attempts to demonstrate the invariance by showing how the shapes and or sizes of the triangles, squares and rectangles can be changed without changing the structural relationships.
This was inspired by a demonstration originally provided by Norman Foo, using different transformations:
http://www.cse.unsw.edu.au/~norman/Pythag.html
One of the striking facts about Pythagoras' theorem is how many different ways it can be, and has been, proved.

NB: The programs that present such proofs do not themselves understand the proofs. They can be powerful "cognitive prosthetics" for humans learning mathematics, but the programs do not know what they have done, or why they have done it, and do not understand the invariants involved -- e.g. essentially the same proof could have started with a triangle with different angles, or a triangle of a different size.

I'll now return to the consideration of areas of triangles and how the area of a triangle is altered by moving one vertex, extending the ideas used in discussing the Median Stretch Theorem and Side Stretch Theorem, above.

BACK TO CONTENTS

Added 9 Feb 2013 - Modified 4 Mar 2013: Another proof of the sum theorem, by Kay Hughes

Last night I was talking about education with Kay Hughes and asked if she could remember how to prove the Triangle Sum Theorem. She could not remember a proof but quickly thought up a proof which I had never previously encountered. My presentation here does not use her words, but offers a more explicit elaboration of the ideas she presented. The key idea was to use a theorem about the sum of external angles of a polygon always being a whole rotation (360 degrees) and combining that with the fact that each such external angle has an internal angle as complement, as explained in more detail below.

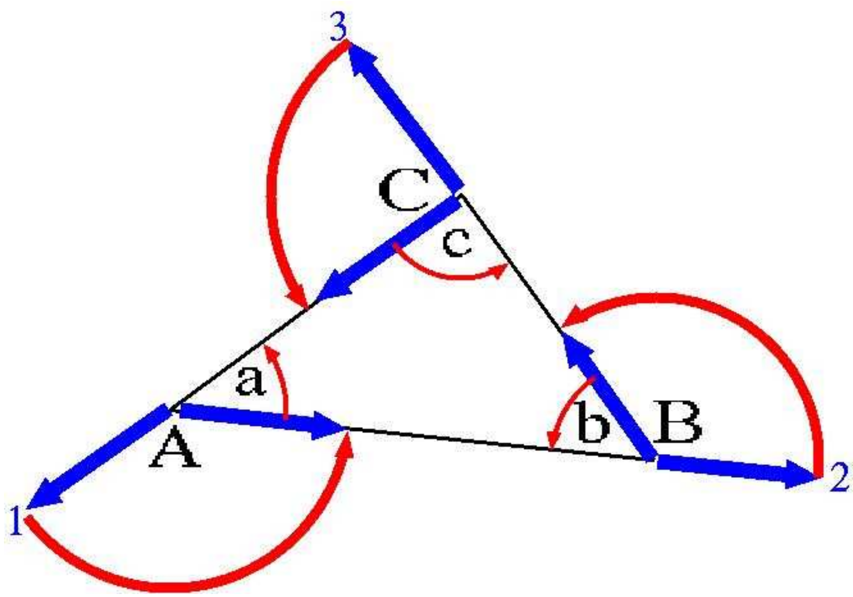


Figure Ang4:

It should be obvious from the figure that it presents a proof that the exterior anti-clockwise angles of a triangle sum to a circle (360 degrees) as do the exterior **clockwise** angles, not shown in the figure.
Added 19 Mar 2013: This was named "The total turtle trip theorem" by Seymour Papert, in his **Mindstorms: Children, Computers, and Powerful Ideas** (1978), though it was well known long before then. (It can be generalised to smooth simple closed curves. See also http://en.wikipedia.org/wiki/Total_curvature .)
The exterior anti-clockwise angles are those obtained by extending each side in turn in one direction then rotating the extension to line up with the next side. So, for example, in Figure Ang4, the internal angles are **a**, **b** and **c** and the exterior anti-clockwise angles **A**, **B** and **C** are got by extending the first side to location **1** then rotating the extension through angle **A** to the next side, then extending that side to location **2** and rotating the extension through angle **B** to the second side, and so on.

Because results of all those rotations bring the rotated arrows back to the original orientation, indicated at **1** in the figure, and the rotated arrow does not pass through its original direction, the total external anti-clockwise rotation must be a full circle (i.e. 360 degrees). An exercise left to the reader is to show that that's true not only for triangles but for all polygons, and, by symmetry, must also be true for the sum of the clockwise external angles. So:
Theorem External: A + B + C = 360

But each of the internal angles is the **complement** of the adjacent internal angle, because they sum to a straight line. So we have these three truths:
Theorem: A + a = 180 therefore **a = 180 - A**
Theorem: B + b = 180 therefore **b = 180 - A**
Theorem: C + c = 180 therefore **c = 180 - A**
So, the sum of the internal angles is
a + b + c = (180 - A) + (180 - B) + (180 - C)
= 180 + (180 + 180) - (A + B + C)
= 180 + 360 - (A + B + C)
Then substituting from **Theorem External:**
= 180 + 360 - 360
= 180
So, we have another proof of the standard Triangle Sum Theorem:
Theorem Internal: a + b + c = 180

I tried searching for that proof using google and did not find a previous occurrence of it, though there seem to be many web sites that mention both the triangle sum theorem for interior angles and the theorem about exterior angles always summing to 360.

BACK TO CONTENTS

The perpendicular stretch theorem (The need for case analysis in some proofs)

The Median Stretch Theorem (MST above), and the Side Stretch Theorem (SST) on which it depends, both require a single diagram. Distortions of the diagram may produce new figures that look different but they do not require any new form of reasoning.

However there are some theorems in Euclidean geometry whose proof requires use of more than one diagram, because the theorem has a kind of generality that covers structurally different cases. An example of such a theorem is a proof that the area of a triangle is half the area of a rectangle with the same base length and the same height: Area = 0.5 x Base x Height. The reason for requiring more than one diagram (unless there is a proof I have not encountered) will be explained below.

A non-diagrammatic algebraic proof may be possible using the Cartesian-coordinate based representation of geometry, but that is not what this discussion is about.

It is highly regrettable that our educational system produces many people who have simply memorised the Area formula, without ever discovering a proof or being shown one, or even being told that there is a proof, though some may have done experiments weighing triangular and rectangular cards. I shall try to explain how this formula could be proved, though I'll expand the usual proof to help bring out differences between this theorem and previous theorems, explaining why this theorem requires different cases to be dealt with differently.

Consider a theorem related to Figure P-a below, which is subtly different from Figure M (a), above.

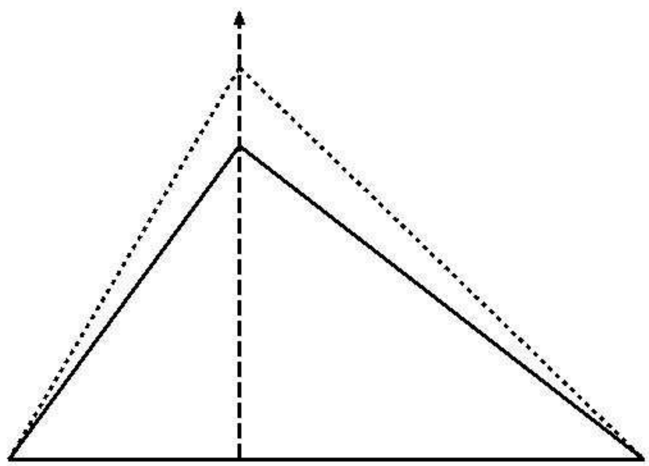


Figure P-a:

Figure P-a includes a straight line drawn between a vertex of the triangle and the opposite side, extended beyond the vertex as indicated by the dashed arrow. In figure M the line used was a **median**, joining the mid-point of a side to the opposite vertex. Here the line is not a median but is **perpendicular** to the opposite side. (In some cases the median and the perpendicular are the same line. Which cases?)

You should find it **obvious** that if the top vertex of the triangle with solid black sides shown in Figure P-a, above, is moved further away from the opposite side (the base), along a line perpendicular to the opposite side (the dashed arrow), then the area enclosed by the triangle must increase. This could be called the "Perpendicular Stretch Theorem" (PST), in contrast with the "Median Stretch Theorem" (MST), which used a line drawn from the middle of the base.

In this figure it is obvious that moving the vertex up the perpendicular will produce a new triangle that encloses the original one. Figure P-a shows why it is obvious, though the Side Stretch Theorem shown in figure S, above, could used to prove this, by dividing the figure into two parts, just as it was used to prove the Median stretch theorem. (As with MST, there is a corresponding theorem about the area decreasing if the vertex moves in the opposite direction on the perpendicular.)

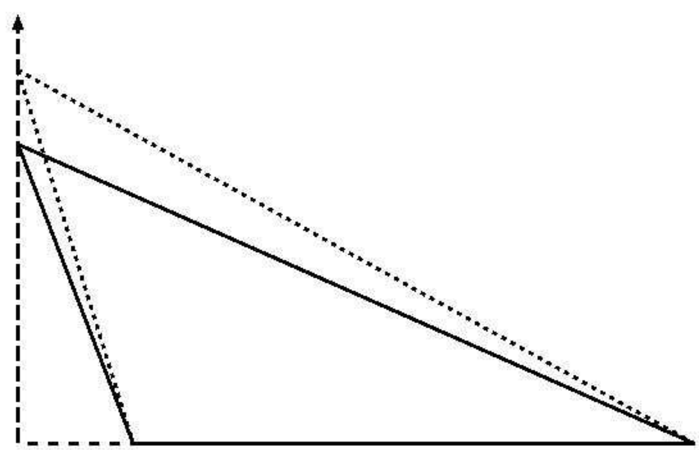
But there is a problem, which you may have noticed, a problem that did not arise for the median stretch theorem. The problem is that whereas any median from the midpoint of one side to the opposite vertex will go through the interior of the triangle, the perpendicular from a side to the opposite vertex may not go through the interior of the triangle, a problem portended by part (b) of Figure M.

BACK TO CONTENTS

A problem with the proof using Figure P-a

Observant readers may have noticed that the reasoning based on Figure P-a has a flaw, since not all movements of a vertex of a triangle perpendicularly away from the opposite side will produce a new triangle that encloses the original one: for example if one of the interior angles (e.g. the one in the left in Figure P-b, below) is obtuse (greater than a right angle), so that the top vertex does not start off perpendicularly above the base of the triangle. The line perpendicular to the "base" that passes through the vertex need not pass through the base, though it will pass through a larger line extending the base, as shown in Figure P-b, which is derived from Figure P-a, by shifting the upper vertex over to the left, so that the perpendicular indicated by the dotted arrow moves outside the triangle, and no longer intersects the base (the side opposite the vertex under consideration), though it intersects the line extending the base.

Figure P-b:



In this case, moving the top vertex upwards will not produce a new triangle enclosing in the old one, because one of the sides of the triangle will move so as to cross the triangle, as illustrated in Figure P-b. So now the proof that the area increases cannot be based on containment: the new triangle produced by moving the vertex upward does not include the old triangle, as in the previous configuration. Is there a way of reasoning about this new configuration so as to demonstrate an invariant relation between direction of motion of the vertex and whether the area of the triangle increases or decreases?

Some readers may notice a way of modifying the proof to deal with figure P-b, thereby extending the proof that moving the vertex further from the line in which the opposite side lies, always increases the area. It is an extension insofar as it covers more cases. Of course, the original proof covered an infinite set of cases, but that infinite set can be extended.

A clue as to how to proceed can come from considering how to prove that moving the vertex of a triangle parallel to the opposite side, as illustrated in Figure P-c, below, cannot change the area.

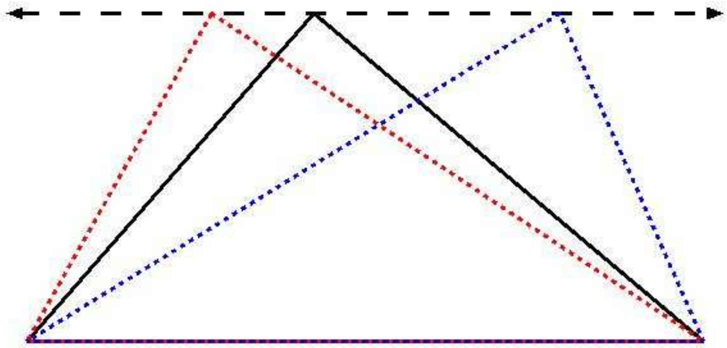


Figure P-c:

TO BE CONTINUED

I shall later extend this discussion by showing how to relate the area of a triangle to the area of a rectangle enclosing it. It will turn out that the triangle must always have half the area of the rectangle, if the rectangle has one side equal in length to a side of the triangle and the other side equal in length to the perpendicular of the triangle. Proving this requires dealing with figures P-a and P-b separately.

The proof using a rectangle requires introducing a new discontinuity into the configuration: dividing up regions of the plane so that they can be compared, added, and subtracted.

Some readers will be tempted to prove the result by using a standard formula for the area of a triangle. In that case they first need to prove that the formula covers all cases, including the sort of triangle shown in Figure P-b.

For anyone interested, here’s a hint. Consider Figure TriRect, below. Try to prove that every triangle can be given an enclosing rectangle, such that every vertex of the triangle is on a side of the rectangle and two of the vertices are on one side of the rectangle, and at least two of the vertices of the triangle lie on vertices of the rectangle.

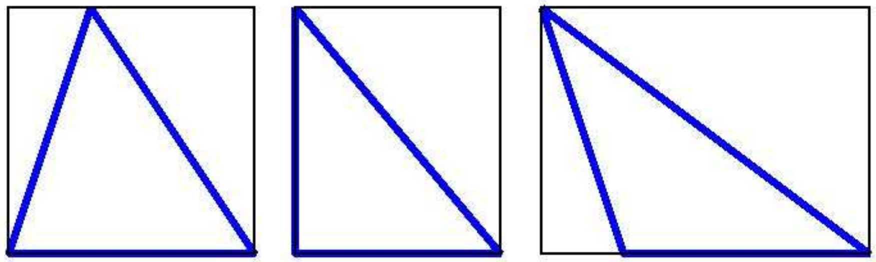


Figure TriRect:

Can you prove something about the area of a triangle by considering such enclosing rectangles?

Many mathematical proofs are concerned with cases that differ in ways that require different proofs, though sometimes there is a way of re-formulating the proof so that the same reasoning applies to all the structurally distinct cases. A fascinating series of examples from the history of mathematics is presented in (Lakatos, 1976)

> A hard problem for human and animal psychology, and studies of evolution of cognition, is to explain how humans (and presumably some other animals capable of intelligent reasoning about their affordances), are able to perform these feats. How do their brains, or their minds (the virtual information-processing machines running on their brains), become aware that the special case being perceived shares structure and consequences of that structure, with infinitely many other configurations, the majority of which have never before been seen or thought about.
>
> For an explanation of the notion of a virtual machine made of other concurrently active virtual machines, some of which also interact with the environment, see http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html

BACK TO CONTENTS

Toward robot mathematicians discovering geometry

It will be some time before we have robot mathematicians that understand Pardoe's proof, or the proofs of the 'Stretch' theorems summarised above (Median stretch, Side stretch, Perpendicular stretch theorems), or can think about how to compute the area of a triangle, or can discover the existence of prime numbers by playing with blocks (in the manner described here), or can perceive and make use of the many different sorts of affordance that humans and other animals can cope with (including, in the case of humans: proto-affordances, action affordances, vicarious affordances, epistemic affordances, deliberative affordances, communicative affordances), many described in this presentation on Gibson's theories.

Even longer before a robot mathematician spontaneously re-invents Pardoe's proof? (Or the proofs in Nelsen's book.)

For some speculations about evolution of mathematical competences see

- http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#mathcog
  If learning maths requires a teacher, where did the first teachers come from?
- http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#toddler

Why (and how) did biological evolution produce mathematicians?

- http://www.cs.bham.ac.uk/research/projects/cosy/papers#tr0802
  Kantian Philosophy of Mathematics and Young Robots (MKM08)
- http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0807
  The Well-Designed Young Mathematician (AI Journal 2008)

Chemical computation

A deeper question is whether there is something about the information-processing engines developed and used by evolution that are not modelled in turing machines or modern computing systems, or have totally intractable complexity on Turing machines or modern computers. I shall later produce some speculative notes on whether there are deep differences between chemistry-based computation and more familiar forms of computation.

If there are differences I suspect they may depend on some of the following:

- Chemical processes involve both continuous changes in spatial and structural relations and also the ability to cross a phase boundary and snap into (or out of) a discrete stable state that resists change by thermal buffeting and other processes.
  This stability could rely on quantum mechanisms.
- They also allow multiple constraints to be exercised by complex wholes on parts, which allow certain forms of motion or rotation or chemical behaviours but not others.
- Some switches between discrete states, or between fixed and continuously variable states can be controlled at low cost in energy by catalytic mechanisms.

It is clear that organisms used chemical computation long before neural or other forms were available. Even in organisms with brains, chemical information processing persists and plays a more fundamental role (e.g. building brains and supporting their functionality). This is just a question: I have no answers at present, but watch this space, and this PDF slide presentation on Meta-Morphogenesis (still work in progress):
http://tinyurl.com/CogTalks#talk107

BACK TO CONTENTS

Comparison with logical proofs

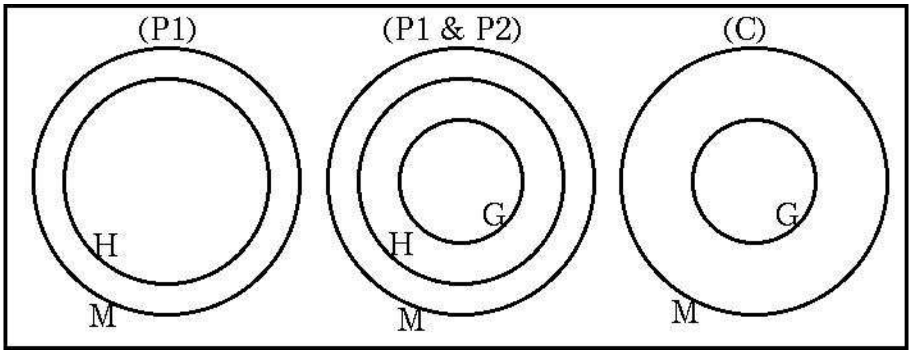
Many mathematical proofs involve sequences of logical formulae or equations, with something altered between stages in the sequence. Those sequences can be thought of as processes, but they are essentially discrete, discontinuous processes. For example, consider the transformation from (P1) and (P2) to (C) in this logical proof:

Premisses
(P1) All Humans are Mortal (or (All x)H(x) -> M(x))
AND
(P2) All Greeks are Humans (or (All x)G(x) -> H(x))

Conclusion
(C) All Greeks are Mortal (or (All x)G(x) -> M(x))

For someone who does not find this obvious, the proof can first be transformed into a diagram which initially represents (P1), then adds the information in (P2), then shows how that includes the information in (C), showing the proof to be valid.

This can be thought of as a process, but the steps are distinct and there are not meaningful intermediate stages, e.g in which the antecedent "H(x)" and the implication arrow "->" are gradually removed from the original implication, and the word "Socrates" gradually replaces the variable "x". Nevertheless the proof can be expressed diagrammatically using Euler Circles as in Figure Syll (often confused with Venn Diagrams, which could also be used).

Figure Syll: 

In (Sloman, 1971) I suggested that both types of proof could be regarded as involving operations on representations that are guaranteed to "preserve denotation". This is an oversimplification, but perhaps an extension of that idea can be made to work.

In the Pardoe proof, "preserving denotation" would have to imply that a process starting with the initial configuration in Figure Ang3, and keeping the triangle unchanged throughout, could go through the stages in the successive configurations depicted, without anything in the state of affairs being depicted changing to accommodate the depiction, apart from the changes in position and orientation of the arrow, as depicted. This implies, for example, that there are no damaging operations on the material of which the structures are composed. (I suspect there is a better way to express all this.)

Cathy Legg has presented some of the ideas of C.S. Peirce on diagrammatic reasoning in (Legg 2011) It is not clear to me whether Peirce's ideas can be usefully applied to the kinds of reasoning discussed here, which are concerned with geometrical reasoning as a biological phenomenon with roots in pre-human cognition, and properties that I suspect could be replicated in robots, but have not yet been, in part because the phenomena have not yet been understood.

BACK TO CONTENTS

A partial list of references (to be expanded)

- http://math.berkeley.edu/~rbayer/09su-55/handouts/ProofByPicture-printable.pdf
  Robertson Bayer, Proof By Picture (PDF lecture slides),
  University of California, Berkeley Math 55, Summer 2009
  (A collection of diagrammatic proofs of mathematical theorems, most of them non-geometric -- e.g. geometric proofs of theorems in number theory.
  Includes the 'Chinese' proof of Pythagoras' Theorem.)

- Kenneth Craik, **The Nature of Explanation,**
  Cambridge University Press, 1943, London, New York,

  NOTE:
  Craik proposed that biological evolution produced animals with the ability to work out what the consequences of an action would be without performing the action, by making use of an abstract **model** of the situation in which the action is performed. It is not clear to me that he noticed the difference between running a detailed model of a specific situation to discover the specific consequences, which some current AI systems (e.g. game-engines) can do, and noticing an invariant property of such a process with different starting configurations as required for understanding why a strategy will work in a (possibly infinite) class of cases.
  I think he came close, but did not quite get there, but I have read only the 1943 book.

- Susan Blackmore, **Zen and the Art of Consciousness,** Oneworld Publications, 2011
  (Previous title: Ten Zen Questions,)
  Discussed in:
  http://www.cs.bham.ac.uk/research/projects/cogaff/misc/blackmore-zen-consciousness.html

- George B. Dyson,
  Darwin Among The Machines: The Evolution Of Global Intelligence,
  Addison-Wesley, 1997,

- **Entertaining video tutorial on Pythagoras and irrationality**
  showing the connection between irrationality of square root of 2, and Pythagoras' theorem.
  https://www.youtube.com/watch?v=X1E7I7_r3Cw
  "What's up with Pythagoras", by Vi Hart. See more of her work:
  https://www.youtube.com/user/Vihart
  And here: http://vihart.com
  You may have to pause and replay bits of her videos, to take in some of the details.

- Immanuel Kant, **Critique of Pure Reason,** 1781,
  Translated (1929) by Norman Kemp Smith, London, Macmillan,

- Annette Karmiloff-Smith,
  Beyond Modularity: A Developmental Perspective on Cognitive Science,
  MIT Press, 1992, Summarised in: http://tinyurl.com/CogMisc/beyond-modularity.html

- http://adrenaline.ucsd.edu/Kirsh/Articles/Earwig/earwig-cleaned.html
  David Kirsh,
  "Today the earwig, tomorrow man?,
  in **Artificial Intelligence,** 47, 1, pp. 161--184, 1991

- Imre Lakatos, **Proofs and Refutations,**
  1976, CUP, Cambridge, UK,

- Roger B. Nelsen, (Often mis-spellt as "Nelson")
  Proofs without words: Exercises in Visual Thinking,
  Mathematical Association of America, Washington DC, 1993,

- http://hdl.handle.net/10289/5153
  Catherine Legg, What is a logical diagram?
  Paper presented at Mini-Conference on Logical Pragmatism,

Auckland, New Zealand, February 25, 2011.

- **Mary Leng interviews Mateja Jamnik on Spatial Reasoning**,
  In **The Reasoner** 7(1), Jan 2012. pp 1--4.
  http://www.kent.ac.uk/secl/philosophy/jw/TheReasoner/vol7/TheReasoner-7(1).pdf

- J. Sauvy and S. Sauvy, **The Child's Discovery of Space: From hopscotch to mazes -- an introduction to intuitive topology,**
  Penguin Education, 1974, Translated from the French by Pam Wells,

- http://tinyurl.com/BhamCog/04.html#200407
  Aaron Sloman, 'Interactions between philosophy and AI:
  The role of intuition and non-logical reasoning in intelligence',
  Proc 2nd IJCAI, 1971, London, pp. 209--226, William Kaufmann,
  Also reprinted in AI journal, 1971, and in "The Computer Revolution in Philosophy", Chapter 7, 1978)
  http://tinyurl.com/BhamCog/crp/chap7.html
  That article was originally written as a critique of McCarthy and Hayes 1969:
  "Some philosophical problems from the standpoint of Artificial Intelligence"
  (In **Machine Intelligence 4**)
- http://tinyurl.com/BhamCog/96-99.html#15
  Aaron Sloman, Actual Possibilities, in
  Principles of Knowledge Representation and Reasoning: Proc. 5th Int. Conf. (KR '96),
  Eds. L.C. Aiello and S.C. Shapiro, Morgan Kaufmann, Boston, MA, 1996, pp. 627--638,

- http://dspace.mit.edu/handle/1721.1/6894
  Gerald. J. Sussman,
  A computational model of skill acquisition, American Elsevier, 1975,

- Max Wertheimer, **Productive Thinking**
  New York, Harper and Brothers, 1945

Acknowledgements

See the acknowledgements section of the paper on P-Geometry
http://tinyurl.com/CogMisc/p-geometry.html#acknowledge

Offers of help in making progress will be accepted gratefully, especially suggestions regarding mechanisms that could enable robots to have an intuitive understanding of space and time that would enable some of them to rediscover Euclidean geometry, including Mary Pardoe's proof.

I believe that could turn out to be a deep vindication of Immanuel Kant's philosophy of mathematics. Some initial thoughts are in my online talks, including

```
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#toddler
Why (and how) did biological evolution produce mathematicians?
```

Maintained by
Aaron Sloman
School of Computer Science
The University of Birmingham

-