Proofs and stories

Family resemblances and family history

Apostolos Doxiadis

June 2007

Draft copy

Do not quote or circulate without the author's permission

CONTENTS

1. Introduction 1
2. The rationale for a family resemblance 4
3. Stories 16
3.1 A preamble: on goals 16
3.2 Action and choice 18
3.3 Are stories really linear? 22
3.4 Stories as webs of photos 24
3.5 What's in a movie: glue, time and causality 26
3.6 Linearization: necessary evil *and* blessing in disguise 38
3.7 Outlines 41
4. Proofs 49
4.1 The curious incident of the proof in the nighttime 50
4.2 Photos and movies in proofs 51
4.3 On links in P-space and the nature of deductive movies 54
4.4 Proof-as-discovered, proof-as-published 65
4.5 Seeing, plans and goals: moving around in P-space 68
4.6 Chance, necessity and structure 73
4.7 What's in a plan? 78
4.8 Similarities of stories and proofs, via the photos-and-movies language 82
5. …Two stories in photos and movies 86
5.1 The old story 86
5.2 The *very* old story 101

1. Introduction

In the spring of the year 2000, a reader's online comment on my novel *Uncle Petros and Goldbach's Conjecture* sparked-off a sequence of thoughts, whose present state of development is recorded in this paper. The comment was: "The book unfolds much as solving a mathematical problem would." [1] (Though it may not sound that way to some, this was meant as a compliment.) At first, I thought the comment to be true in a rather trivial sense: as the novel recounts the attempts of its hero to prove a famous conjecture, and the story is about the complications related to these attempts, it quite naturally "unfolds" like the solution of a problem, or rather like the attempts at it – incidentally, I guess "solution of a problem" is how a non-mathematical person would describe, simply, *proof*.

Yet, when I was asked, a few months later, to represent the side of fiction in a "Mathematics and Culture" conference in Venice, the reader's comment came back to my mind and I decided, mostly for argument's sake, to see whether a story unfolding like a proof could be spoken of in a more general sense, i.e. whether a story which is *not* about the proof of a mathematical theorem – and I hear there are some of those still around – could be said to be proof-like.

Being by trade a novelist and not a scientist, and thus a firm upholder of a *se non è vero è ben trovato* kind of epistemology, I tried to build a case that stories indeed *are* proof-like – and also, along the way, that proofs are story-like. In this, I was perhaps also inspired by the old Greek teachers of oratory, who liked to take the most extreme positions and try and build convincing arguments for them. In fact, in adopting the role of the sophist, I was playing devil's advocate to Socrates' position, expressed in Book Ten of Plato's *Republic*, that there is a *palaia diaphora* (an 'old difference' or 'disagreement') between philosophy and poetry. Supported by his inclusion of Homer and the dramatists in the ranks of the poets, I saw this *diaphora* extending to the ever-present rift between storytelling and mathematics, which rather prototypically embodies C. P. Snow's "two cultures" dichotomy.

Of course, it's extremely easy to list arguments in support of this rift, things which show that storytelling is from Venus and mathematics from Mars. Received wisdom has it that storytelling is addressed to the heart, mathematics to the brain; storytelling has characters, mathematics does not – well, at least not in the usual sense; storytelling is about recognizably human situations and emotions, mathematics is about numbers, shapes, equations, and so on; storytelling describes worlds which are concrete, worlds which even when they are not 'real' – as they are not in mythology or science fiction

---

[1] Mishka M. Mourani in an Amazon.com "customer review" of *Uncle Petros*.

–, are very much understandable to human beings: Polyphemus the Cyclops, HAL the Computer and Winnie the Pooh may not be exactly human, but they behave in ways which are comprehensible to us by virtue of their being anthropomorphic – not so "differential manifolds", "modular forms", "cohomology" and suchlike, the stars of mathematics. Nobody but professional mathematicians – and often even not they, when the concepts are outside their particular, usually narrow, field of specialization – knows what the blessed things are.

Alright, there are great differences, granted. But what about similarities? The first thing I did for my Venice talk, was to turn from comparing the worlds of fiction and mathematics in the abstract, to try and look at the activity of the people who were central to them. And that brought me to comparing *stories* and *proofs* – for creating these is what storytellers and mathematicians do all day. To my surprise, I found these processes to be much more similar than I had thought. But exposing this similarity, which in my talk I called rather grandly an *isomorphism*, depended on my creating for both a particular language, the rationale of which was a bit too formalistic for my taste. To ground this language in some kind of lived experience, I tried to go one step further, in fact to create a *triad* of isomorphisms, also connecting storytelling and proof, through the transitive quality, via a third process: spatial exploration.

Though this may seem like solving a problem by creating another, more complex one, the inclusion of the spatial element was what made the central argument begin to make some cognitive sense – after all, I had started to speak of the similarity of proofs to stories thinking of both of them as quests, and the quest metaphor definitely rests on the hard ground of space. Since then, I returned again and again to the idea of the structural similarity of proofs to stories, in a series of lectures the last of which was at the "Mathematics and Narrative" meeting that took place on Mykonos, in July 2005[2]. And it is to this idea that I come once again, now hoping to be able to make a better case for it in ways which are more natural, i.e. related to the form and function of human cognition. I will describe both stories and proofs in what I shall call a *photos-and-movies language* – the rationale of the term will be explained –, which has the advantage that it is based on a strong underlying cognitive logic, the same logic, in fact, which connects, in historical order, spatial exploration to stories, and stories to proofs.

---

[2] The full text of my Venice talk, titled "Euclid's Poetics", is on my website *www.apostolosdoxiadis.com* in Proofs and Stories. The argument is further developed in four papers, also on my website: "Embedding mathematics in the soul: narrative as a tool in mathematics education", "The Mystery of the Black Knight's Noetherian Ring", "The mathematical logic of narrative", and "Narrative as knowledge" – the last one is unfortunately in Greek, but a much longer English version will soon be posted. The Mykonos lecture was in slideshow and has not been written up.

Section 2 of this paper is a sketch of the argument connecting proofs to stories, based on the concept of the quest – it is this that forms the basis for the photos-and-movies language. I outline here the central reason why I believe that a structural similarity of stories and proofs makes strong cognitive sense, before I go into the description of the language.

In Sections 3 and 4, I will describe and investigate various aspects of the photos-and-movies language, for stories and proofs, respectively. Section 3 deals with stories, and the ways in which the photos-and-movies language shows forth dimensions of narrative that have a strong cognitive side. Then, in Section 4, I try and show how this language applies quite naturally also to proofs. Having done some of the basic exposition of the photos-and-movies language in the previous section, I try to show how most of what we saw of it as applied to stories, is also true of proofs. I end Section 4 with a list of the most basic similarities of stories and proofs, as revealed by this language.

In Section 5, I will present – alas, much more briefly than they deserve – two historical applications. The first has to do with how the photos-and-movies language helps us understand the genesis of logico-deductive mathematics in 5th century BCE Greece, and the way in which proof may in fact have sprang out of a narrative infrastructure. The second refers to the Upper Paleolithic era, and provides a short scenario of how the structural similarity of stories to spatial quests may also account for the birth of narrative. This second section offers no more than a basic indication of the central ideas, which I plan to develop at greater length elsewhere.

Obviously, the arguments I put forth, especially in section 5, are not just inter- but multi-disciplinary. To try and convince you that I have a good command of narratology *and* mathematics *and* cognitive science *and* the history of early science *and* evolutionary psychology *and* anthropology, and a lot more besides, would be to try to convince you that I am a conceited fool. But I am comforted by the fact that there is probably no one on this planet who combines these dimensions of expertise: one lifetime is far too short to acquire them.

That's the good thing about exploring terra incognita, after all: you may not know exactly where you are going – but no one can prove it, because no one has a map.

2. The rationale for a family resemblance

My basic argument develops from two propositions which are more or less easily demonstrable, as well as a far stranger one.

The first two are:

Proposition 1. The quest for a mathematical proof is, precisely, a *quest*.

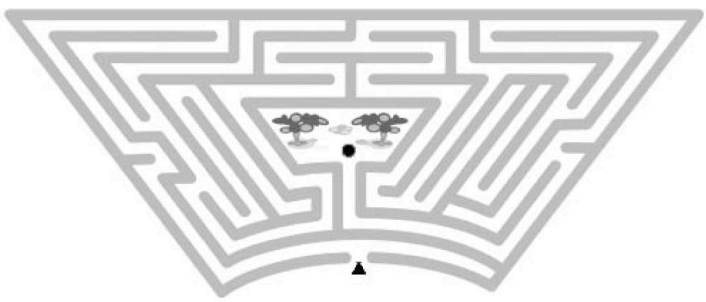
Proposition 2. The structure of a quest story mostly follows the structure of the underlying quest.

These will be discussed at some length in what follows, so if there are any disbelievers, let them please wait. The third proposition, the extravagant one, is a variation of something I read in a popular book on screenwriting, something which would not leave me alone ever since. The exact quote, and discussion, appears further down. For now, I will rephrase it thus:

Proposition 3. All stories are essentially quest stories.

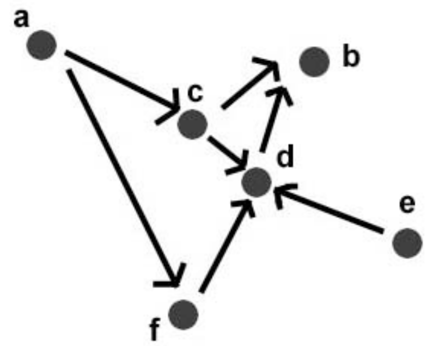
But let's leave Proposition 3 aside for a while and look at the less problematic Propositions 1-2. In both, the dominant word is *quest*: proofs are *quests*, *stories of quests* look like *quests*... Speaking of quests brings naturally to mind the spatial analogy, for space is the original environment for quests.

Here is a well known setting for a quest, the famous maze at Hampton Court Palace:



In this diagram we see all the basic ingredients for a quest, whether it be spatial or not: a) a *searcher*, (here this can be you if you want), placed at b) an *initial position* (the little triangle), aiming for c) the *destination* or *goal*, i.e. that which is searched-for, here represented by the dot, with everything occurring in, d) an *environment* (the actual maze), preferably complex enough to make the search non-trivial.

Now, though this diagram is quite helpful, quests are easier to think about mathematically in the more operational language of what mathematicians call *graphs*. These consist of *dots* and *arrows* or – if you don't have restrictions about some arrows being one-way – little *lines*. This is an example of a graph[3]:



A graph can perfectly represent the environment of a quest, with each dot being a crossroads or, more generally, a point where a decision has to be made, whether to go one way or another – of course some dots, in some graphs, like dot *e* in the above example, don't leave but one option –, and the arrows being the roads which you can take, from dot to dot.

In a quest situation translated into graph language, the graph is your *environment*; your *initial position* is a particular dot on the graph, and your *destination* another one. Your task as a *searcher* is to find a *path*, i.e. a connected sequence of arrows, end-of-one-to-beginning-of-next, leading from the initial position to the destination and moving only in the direction indicated each time by the arrows.

Now, let's look at Proposition 1, above, which says that the attempt to prove a theorem is really a quest. What this means is that even though what the old Greek mathematicians called *porisms* – i.e. truths arrived-at incidentally, in the process of the search for some other truth – do exist in mathematics, most mathematical research is motivated by attempts to prove this or that theorem, or solve this or that problem through conscious, goal-oriented effort. I think that this statement, with the caveat about *porisms*, would be accepted by most working mathematicians.

Like any quest, a search for the proof of a theorem can be thought of as occurring in the more formal environment of a graph, which is actually the

[3] A non-mathematical friend reading an earlier form of this draft was confused, as he said that he always thought that a graph was what he what was called at school "the graph of a curve", i.e. a graphic representation of a function on the x-y axis. Lest this word confuse anyone else, I make it clear that the word "graph" in this paper only refers to objects like the one on this page, the diagrams of dots and arrows which form the object of study of *graph theory*.

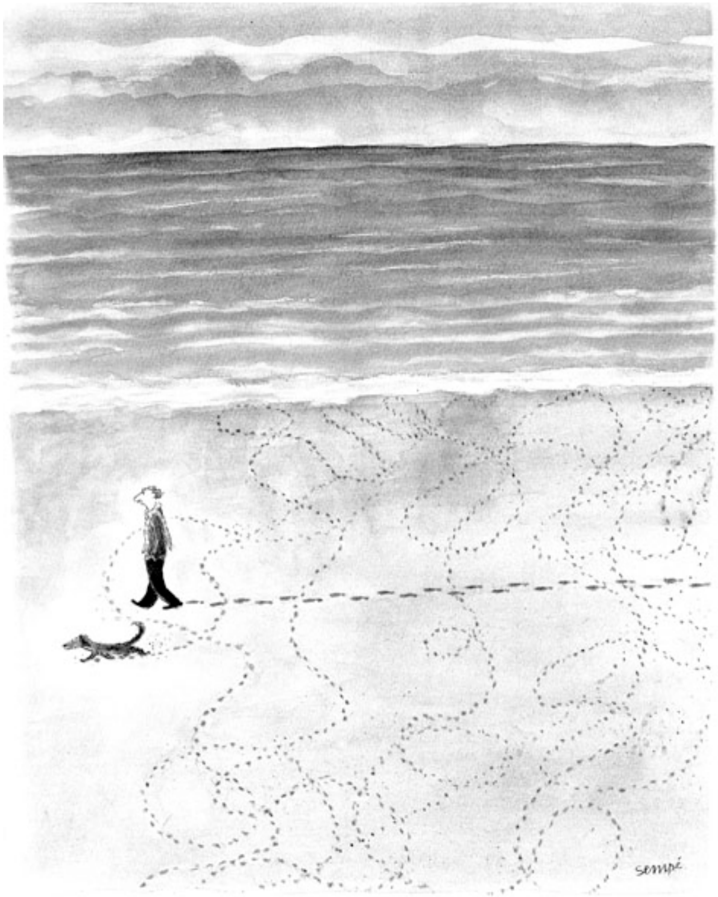
subgraph of a much much bigger one, that of the branch of mathematics to which the particular theorem belongs. The dots of the graph in this case will represent mathematical statements of fact. One could be, for example, that "7 is a prime number" and another, more general one, the statement of the Pythagorean Theorem. But we must also have statements like "9 is a prime number", which is wrong, as mathematicians often use in their proofs the so-called *indirect proof*, or *reductio ad absurdum*, assuming the false to get to the true. As for the arrows in this graph, they would be the mathematically acceptable forms of connections of statements, i.e. the ways in which you can legitimately go from one to the other, whether these be so-called *rewrite rules* like, for example, being allowed to write the expression $(a+b)^2$ as $a^2 + 2ab + b^2$ if you so wish, or elementary deductions which allow you to jump from fact to fact and saying "therefore" or "thus" or something like that. The searcher (here: the *mathematician*) begins from an initial position, which is the known (here: *axioms* of a theory and/or *already-proven theorems*) and wants to reach a far away destination, a dot in the graph, which is the statement of a *new theorem*.

There is a definition of proof by Alan Turing which defines it as "a long process every step of which is trivial"[4]. Thinking of the universe of all possible propositions of mathematics as a colossally humongous graph – this infinite phantom space started intruding into the consciousness of mathematicians near the end of the 19th century – you can focus at the branch of mathematics which includes the particular theorem you are concerned with and concentrate on a more manageable subgraph of it that will contain its proof: the *environment* of your search. Then – this is what Turing's definition says – you can find inside this subgraph a sort of walk which, step by trivial step, leads from the known to the unknown, and your theorem. This is your proof, which you can then write more or less as a list of the dots you went by.

Of course, a published proof shows the structure of the proof *after* it was discovered, so it does not really show the quest, but a sanitized, cleaned-up version. It is not the adventure of the discovery that it depicts, but a sequence of instructions the explorer has noted, to guide future travelers to the goal in some optimal way. But it's clear – if you doubt it, ask a mathematician! – that the complex subgraph of the space of propositions recording the actual quest will look exceedingly messier. In fact, the proof-as-published and the proof-as-discovered will map tracks that are as dissimilar as the tracks left by the man and the dog in this wonderful drawing by Sempé:

---

[4] I feel very embarrassed at not being able, despite my best efforts, to give you a precise reference for this – or even to assure you that I remember it totally correctly. I read it many years ago in a serious source (= a book!), and despite many efforts and internet quests in the past few years, I have not been able to trace it. If a reader can help me in this in any way, he or she will have my gratitude.



However, both the proof-as-published and the proof-as-discovered can be described and talked-of in graph language – though of course, each one will result in a different graph.

Talking of proofs-as-discovered, we must remember the important part played by statements that are not yet – and may never be – proven, what mathematicians call *hypotheses*, or *conjectures*[5]. In graph theory language, this would mean that you can set a certain dot as your destination, and then attempt to find whether there exists a path to it and, if so, which one it is. This introduces a fundamental distinction with a strong cognitive analogue, that between "seeing" and "going to" a certain place: we can "see" a conjecture, in the sense of being able to state the proposition, but if we want to prove it we have to "go" to it, i.e. construct a course that leads from the already-known to the unknown, the conjecture.

[5] For an interesting discussion, see Barry Mazur's article "Conjecture", *Synthèse*, Volume 111, Number 2, May 1997.

Now, look at the next image. We all know the difference in difficulty between a viewer of this landscape in real space *seeing* the peak of the mountain in the distance from where s/he is standing, and actually *going there*. Add to that the possibility that the peak may be a mirage – since a conjecture may not turn out to be true – and you have a colorful metaphor for mathematical research.

Clearly, to speak of a *theorem*, instead of a conjecture, the "going to" must have been completed, in a mathematically proper way. But were it not for "seeing", i.e. for guessing at propositions that you cannot still connect too, mathematics would not exist.

To put it another way: if you are moving around in a substantial subgraph of the humongous graph of a theory, or even of the whole of mathematics, "seeing" becomes essential, for it gives you a *goal*.

Viewing this again in spatial terms: the area that you command as certain knowledge – i.e. the type of knowledge which mathematics is famous for possessing – is the already-proven[6]. Now, in order to venture from the known to the unknown, in such an environment, you have two ways: a) to perform random combinatorial games with the known, in the hope of coming up with some interesting new results and b) to search towards specific destinations. Of course, the second method will dramatically decrease your options. But it has the advantage that it will ward-off the beasts lurking in immense complexity and thus will, equally dramatically, increase your

[6] In Barry Mazur's aforementioned paper, he recounts how a Dutch friend enlightened him on the true essence of mathematics by giving him the Dutch word for it: *wiskunde*, a composite of *wis*, which means *sure*, and *kunde*, meaning *science* or *domain of knowledge*. Thus, Barry says, mathematics is '*sureology*'.

chances of finding something interesting. The destination, the goal, is precisely the conjecture. As any theorist of complexity will tell you, in even a moderately big structure – and the graph of a whole theory is immoderately big – doing a random, combinatorial search is practically synonymous to doing nothing.

It is enough to accept the translation of a theory into graph language, as described, to accept Proposition 1, i.e. “a proof is a quest”, as true. Our Proposition 2 speaks of the similarity of *quests* to a special kind of stories, i.e. *stories of quests*.

Let’s look first at a more restricted form of it, restricting quest to the special case of *proof*, as indicated by Proposition 1. Proposition 2 now lets us speak of the similarity *of a proof to the story of this proof* – which was more or less the point of the reader’s comment about *Uncle Petros* from which this discussion began.

The quest for a proof is a subgraph of the huge graph of mathematical theories. And if we want to write the story of a proof or, even more so, the story of the attempts at it, then to the extent that we structure the story on the actual process of the quest – and not, say, on the mathematician’s emotional troubles at that time, as did David Auburn in his play *Proof* –, the structure of the story can be drawn as a graph that will look very much like the graph of the proof. To do this in the easiest possible way, take again the dots to represent the statements of fact of the mathematical theory (both true and false), but disguised as events in the mathematician’s life: for example, instead of just saying “statement x”, we’ll say something like, “statement x is now proven”. As for the arrows of the graph, they will still represent the way we get from statement to statement in the mathematical sense, but ever so slightly personalized: so if it’s “factorize” in the proof graph, it will become “our mathematician factorizes” in the story graph.

Small doses of extraneous-to-the-proof human drama will not actually obstruct this structure, not very much anyway, especially as we can mark them, say paint them red so they can be easily identifiable and so extractable and/or ignorable by the more purely mathematically-minded readers of the story of the proof. But as we shall be dealing with the graph of the proof-as-discovered, we should not worry at all about the opportunities for drama: unless the mathematician was incredibly lucky, the story of the proof will certainly contain false starts, cul-de-sacs, failures, loops, repetitions, and so on, i.e. things which you can use in any good story. And you can easily add dramatic flourishes at these points, while recounting the actual quest process, translating for example, “so-and-so’s lemma doesn’t work after all at this point” in the quest graph, to “ ‘our mathematician realized with great despair that so-and-so’s lemma didn’t work after all’ ”. (If you are feeling melodramatic, you can also have your mathematician attempt suicide – this would be a good point for it.)

Based on the above, we could say that if what we wanted to show in this paper was not that "Proofs and stories are really the same thing" but "Proofs and stories of proofs are really the same thing", we would not have too much trouble making a good case of it – always on the condition that we kept our stories of proofs quite clean, with not too much non-proofy stuff in them, or we marked such stuff in ways which would make it ignorable. As for the critical phrase "are really the same thing" in this variation of the theme – in the actual theme too, really – it really means the same as it would if we had written read "Stories and proofs are *very much alike in structure*". Incidentally, the use of "really the same thing" to denote "very much alike in structure" is a habit of modern mathematicians. Even more so, they will sometimes say casually that something *is* something else, actually meaning what a non-mathematician would probably just describe as the other thing: looking very much alike in structure[7]. But as by this new definition of the equality of an x and a y you can still do with x pretty much as you would do with y, and vice versa – i.e. as you would if x and y were really equal in the good old sense –, it's a pretty good definition[8].

So, combining the statement that "proofs <u>are</u> quests" (Proposition 1) and the statement that "proofs <u>are</u> stories of proofs", (Proposition 2) – I underline the "are" to stress that we are speaking of this new sense of equality which is equivalent to "looks structurally the same" –, let's see if we can generalize further, to the proposition that "proofs <u>are</u> stories of quests (quest stories)"[9]. For this to become cognitively meaningful we only have to show that any quest, regardless of the nature of its environment, can be represented in a graph.

This is actually quite easy:

---

[7] A probable ancestor of this tendency, apart from the fact that the human mind thrives on metaphor, could be the use of letters for the description of geometric shapes in Greek geometry, which taught people to call a line on paper by the name 'AB' and to feel that the one *is* the other. And a clear early instance of it can be found in analytic geometry: in its context a student is easily forgiven for saying that the algebraic expression $x^2+y^2=1$ *is* the circle – he/she may be corrected for his/her use of English, but not for his/her mathematics. And from being forgiven for saying something to accepting is as the truth is not too great a step, as can be seen from the birth of the concept of imaginary numbers.

[8] On the matter of this new kind of equality, which mathematicians more formally call *isomorphism*, see also Barry Mazur's "When is one thing equal to some other thing". (Preprint on his site.)

[9] I say here 'quest stories' instead of the synonymous 'stories of quests', also to refer to an accepted category of stories in literature, listing among its member such venerable examples as *Gilgamesh* and the *Odyssey*. It seems very probable that quest stories were among the first stories ever, or even *the* first stories ever, being intimately related to the concerns of the first humans: the hunt for animals and the exploration for new habitations.

To begin with, we can consider the case where the quest is *mathematical* as having been dealt with in the context of Proposition 1, with the dots being mathematical statements and the arrows the legitimate "steps" from one to the other. Now, if the quest in the quest story is *spatial*, as say in a tale of exploration, again we have obvious candidates for our dots and arrows: the dots will be physical locations, and the arrows the journeys from the one to the other.

But what about other cases, where the quest is more *abstract*? Let's take a very non-geographical and non-mathematical case of quest story, i.e. the *whodunit*, preferably of the so-called Golden Age of the form, as defined by the work of writers such as Agatha Christie, Josephine Tey, Rex Stout, and others. A whodunit is indeed a quest story: the *searcher* is the detective; the *initial point* is learning that the crime has been committed; the *destination* is unmasking the culprit, i.e. learning "who done it". Of the four elements of a quest, this leaves only the *environment*. Now, as both origin and destination here represent pieces of information or knowledge, it's reasonable to say that the dots of this quest will be more of the same, i.e. information related to the crime, but again expressed – as we did with proof – as action sentences, since stories have to do with action (thus, not "the cigarette butt", but "the discovery of the cigarette butt").

Like a mathematician, the detective can form conjectures, i.e. can "see" dots (for example, suspecting a particular person of lying) and then try to get evidence for it, i.e. "go" to it. And there may also be many *porisms*, many windfall opportunities for new information. The passages, the arrows of the graph, will be the various methods detectives have for proceeding from piece to piece of knowledge, as for example questioning witnesses, relying on lab work, pure thinking ("little gray cells"), etc. As in a mathematical proof, these arrows will sometimes converge, i.e. the two facts that "the last visitor of the victim was Swedish" combined with the realization that "the cigarette butt is of a Swedish brand" would make the last visitor a strong suspect, worth tailing or calling to the police station for some quality time, or whatever.

If we look at the story's events for their logical content, as opposed to the emotional, and we set aside the detective's family- and/or love-troubles – incidentally, these have become a *sine qua non* of modern mysteries –, the progress of the detective from the discovery of the crime to the revelation of the identity of the culprit can be reasonably well mapped in a graph. As with a mathematical proof, this would not be a mere path but a more complex subgraph, a tree, possibly also including loop paths[10], of the graph of *all possible* pieces of knowledge surrounding the crime, i.e. not just those that

---

[10] Actually, these story graphs will be one step up in complexity from trees, being objects graph theorists call *dags*. For an illustration and a definition, see the bottom of page 40 – but the reader need not do this just yet.

actually turned out to be true, but also the dots referring to the conjectures the detective has to make to eliminate suspects to get, eventually, to the solution – like a mathematician with the *reductio ad absurdum*, the detective will have to use up many wrong hypotheses to arrive eventually at the truth.

Happily for our discussion, the whodunit is more-or-less prototypical of a quest story in an abstract space of possibilities. What makes it so is the clear emphasis on *finding out something*, i.e. on it being an exploration starting from the known and moving towards the unknown, with a clear goal and with the desire for more knowledge as moving force: this is what all quest stories are, if you just switch from the spatial to the abstract context.

So, we can generalize and say that as long as a quest story sticks pretty close to the quest facts we can express it in graph language in the specific way we outlined. Or, even more generally: any story can be quite naturally expressed in the language of graphs, in the way we described, *to the extent that it is a quest story*. And thus now we can say it: to the extent that it is a quest story, any story <u>is</u> (= "looks very much alike, structurally") a proof.

Now, we said that a story looks like a proof "to the extent that it is a quest". But even the most quest-like of stories – and the Golden Age whodunits probably rate as high as fictions can score on this scale, perhaps with the exception of tales of geographical exploration or treasure hunts – have a lot in them that is *not* quest-like. So, if the graph of the structure of the story just results from the graph of the quest, quest-like though the story may be the graph will not really be a very adequate representation of the story, except in the very extreme cases of stories focusing exclusively on the intellectual adventure. But what about the *Odyssey*, perhaps the most famous quest story of all – in which, however, the quest element is not all that strong? Would a graph do any justice to the sophistication of its structure? Or, for that matter, what about all the other stories, those which do not look at all like quests, like the *Antigone*, or *King Lear*, or *Madame Bovary*, or the *The Possessed*? How could these stories, where the heroes are not – at least not in any obvious ways – searchers after a specific truth, translate into the language of graphs in an interesting way?

At this point, we can start to discuss the more extreme Proposition 3, that "all stories are essentially quest stories". Should it be demonstrated – and the purpose of the biggest part of the next section (Section 3) of this paper is devoted precisely to this task, i.e. the construction of graph language for stories – that even the most unquest-like stories in the conventional sense of the term, really *are* quest stories, we will have gone a long way towards generalizing this affinity, and showing the strong proof-like element in any story – or, if you want, the strong story-like elements in any proof.

In a while, we will moderate the scope of the claim of Proposition 3 to a slightly less extreme general principle. But let us begin with the proposition it in its strongest sense[11]. This is the form in which I first read it, as it appears in Robert McKee's book well-known how-to book for writers *Story* (1997), which despite its title is mostly about writing screenplays. I quote McKee at some length:

> …In truth there's only one story. In essence we have told one another the same tale, one way or another, since the dawn of humanity, and the story could be usefully called *the Quest*. All stories take the form of a Quest. For better or worse, an event throws a character's life out of balance, arousing in him the conscious and/or unconscious desire for that which he feels will restore balance, launching him on a Quest for his Object of Desire against forces of antagonism (inner, personal, extra-personal). He may or may not achieve it. This is the story in a nutshell. (*McKee's italics and capitals*.)

As my introducing a how-to book on screenwriting into this kind of discussion may raise some eyebrows, let me pause for a moment to justify my choice. I am an avid reader of manuals on all forms of writing. However, though their authors write and market them as *prescriptive* – i.e., precisely, *how-to* –, I read them as *descriptive*. And if you make that switch, they become eminently more interesting and, I think, deserving of serious attention. It is not merely a question of whether what they say is "right" or not. Their authors, at least of the more successful ones – and success is here of the essence, as it reflects market taste – are not theoreticians of narrative. They are its engineers and marketeers, and as both engineering (i.e. result-oriented, applied knowledge) and marketing (a.k.a. audience-pleasing) are old and hallowed concerns of storytelling, I think that what they have to say is of great interest to anyone concerned with the theoretical side of narrative. After all, where would physics be today were it not for the engineers and marketeers dealing in its ideas?

Actually, description (morphology and taxonomy) and prescription ('how-to') have been going hand in hand for a long long time in the theory of storytelling. In a tradition going back at least, in the Western world, to Aristotle's *Poetics*, which quite definitely has elements of both, even the most practical of how-to books begin by proposing their own version of narrative theory, before they instruct the reader on how to apply it. And though perhaps

---

[11] Though stating a principle in its greatest possible generality, as a rules makes it incorrect in the strict sense, this is often cognitively much more useful than expressing it from the first in a watered-down, but much more realistic version. A good example is Marshall McLuhan's "the medium is the message": though of course what it really means is that "the medium partly affects and/or determines the message", and in this form would be accepted by every reasonable person, it is the extremity of the original statement which made it so powerful, and gave to its later milder forms their staying power.

a more academically-minded theorist would not have been as bold as McKee, in proposing the quest as the underlying structure of *all* stories – that's why progress also needs the down-to-earth, result-orientation of engineers who will bend theory to the needs of practice any day of the week --, the roots of McKee's thesis are in the work of at least two great scholars.

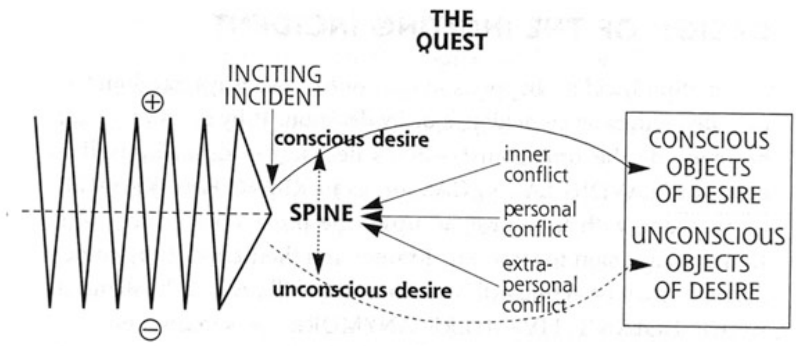
In fact, his presentation of the "all stories are quests" principle, beginning with an event throwing "the character's life out of balance" has a strong resonance of Campbell's *Hero with a Thousand Faces* (1949) and its James Joyce-inspired concept of a basic *monomyth* underlying the mythologies of all cultures[12]. (Campbell's book is listed in McKee's "suggested reading" list.) And of course Campbell's theory of the monomyth – though I don't know if there was a direct influence – has definite structural similarities with Vladimir Propp's "Morphology of the Magical Folktale" (published in 1928 in Russian, but much later in other languages) and his description of the basic structure of the adventure tale, with the initial crisis precipitating the hero's departure. Incidentally, Campbell is also mentioned in Syd Field's *Screenwriting* (1979), which was the "bible" of Hollywood studio script departments for some years, before McKee's *Story*, and then Christopher Vogler's *The Writer's Journey: Mythic structure for writers* (1998) also joined the list. (The last of these declares its purpose quite clearly to be something of a "Joseph Campbell for screenwriters".) [13]

McKee inserts the following image in his book to illustrate his thesis that all quests are stories (p. 197):

---

[12] This idea was born when Campbell was working on his first book, a sort of "user's guide" to *Finnegan's Wake.*

[13] The influence of how-to screenwriting gurus on Hollywood is benignly satirized in Spike Jonze's film *Adaptation*, starring Nicholas Cage. In fact, the actual Robert McKee (though portrayed by an actor) plays a supporting part in the film, and is shown teaching his famous seminars and coaxing innocents, including the screenwriter hero, into conforming to the hallowed rules of the trade, as he sees them. To be fair to McKee, we must emphasize that he is addressing his book to nascent screenwriters, and so his main concern is commercial feature films, which despite any more traditional sub-genre they may belong to, really form a genre of sorts of their own. In this sense, the examples McKee mostly focuses on create a much more uniform sample of stories than if he were looking at the stories of world literature. And, also in this sense, his aphorisms on how a story "should be" make much more sense in the context of profit-oriented, commercial feature films than if read as in a more general context of world fiction.



(Please note that this picture – especially if you look at the structure implied by the curved arrows – portrays a very rough *graph*, rather in the vein that we have been discussing.)

In what follows, I will be more or less inspired by the spirit of McKee's (Campbell's too?) position, that all stories are essentially quest stories, and construct a language by which they can be viewed, analyzed, and depicted from this viewpoint. And it's important to stress here, once again, that the goal of reducing all stories to quests is primarily that of reducing them to graph language, entities in which a quest situation, i.e. having the right *environment* (the graph), and a *searcher*, an *initial position* and a *destination*, can be defined, given the right criteria.

But as stating the principle in its full generality – though it doubtlessly stimulates thought *à la McLuhan* – makes it too blunt an instrument for dissection and analysis, I propose the following modifications and/or caveats. Also, let me add here, that I mostly use the term *story* (rather than *narrative* or *fiction*) to stress the fact that I am referring to pieces of text that would be accepted as narrative in a common sense kind of way, and not stretching my arguments to cover every possible variation of modernist or post-modernist fiction, some of which may be very (and intentionally) un-storylike.

This my amended form of the Proposition 3:

<u>Proposition 3*.</u> All stories have a strong quest dimension. In most of them it is *very* strong, and dominates their structure. Also:

3*.1 The quest dimension of a story applies basically to the elements of plot and character.

3*.2 A quest structure as shown in McKee's graph, sometimes *does not* apply to a story as a whole, though it does to smaller units.

3*.3 The quest dimension can separate a story into substories, either vertically (by characters and action) and/or horizontally (by time).

> 3*.4 The quest structure applies much more strongly to each of the basic substories, at least at some level of analysis.
>
> 3*.5 The substories can interact, compete, clash, complement one another, and so on, in ways which are explicable by the quest dimension.

Now, putting together Propositions 1 and 2 and Proposition 3* we can say that the statement that “proofs are stories” is a direct result of the statement that “stories are quests”, in fact a special case of it.

We shall now go on to discuss in what sense we can talk of a strong quest dimension in all stories, and why, as I believe, the graph language that will naturally support this claim goes to the very essence of storytelling as a cognitive process.

3. Stories

3.1 A preamble: on goals

The rationale of the *photos-and-movies* language is based on the fact that stories evolve in time and are mostly made up of the description of action. But as at the core of time-evolving action sequences are such concepts as *wish*, *intention*, *motivation*, *goal* and *plan*, these will play a big part in our analysis.

Human behavior is not mere information processing nor, except in very few cases, reflex-governed, but at least to some extent planned, i.e. thought-out in advance, in packets of more than one little bit. Our form of more or less forward-looking motivation, combined with an advanced symbolic language, is the shaping power affecting the structure of our behavior, that which makes humans (most probably) the only animals which do not live in a continuous present. Human beings, like all animals, have needs – but, unlike them, they can delay gratification in the service of higher, or sometimes, lower needs. And, more importantly, they can plan farther ahead than any other species, their needs, wishes and intentions being translated into goals that spread out into plans. And though we know what often happens to the best laid plans of men – mice, unless they are named Mickey or Jerry actually *don’t have* plans – that does not stop them being there in the first place, shaping the mega-structure of behavior even when they fail – sometimes especially then.

My belief in the centrality of the dimension of goals and plans to the cognitive understanding of action and narrative comes from different directions. Not counting self-observation, these are:

A. The work of various theorists of human behavior, and more especially two important books, that have been influential in the cognitive sciences: *Plans and the Structure of Behavior* by

George A. Miller, Eugene Galanter and Karl Pribram (1960, NY: Holt, Rinehart and Winston Inc.) and *Scripts, Plans, Goals and Understanding*, by Roger Schank, with Robert Abelson (1977, Hillsdale, NJ: Lawrence Erlbaum and Associates). The second of these also makes a strong connection to the cognitive study of narrative.

B. My work as a theatre-, and film director, and the usefulness in understanding structured action of the ideas of Constantin Stanislavsky, in which an anatomy of human behavior based on motives and what he calls “super objectives” is central. These were later elaborated and applied to the analysis of dramatic writing by another famous how-to book, Lajos Egri’s *The Art of Dramatic Writing: its Basis in the Creative Interpretation of Human Motives* (1946, NY: Touchstone Books).

C. Some rudimentary knowledge of the ideas of computational complexity, and the realization to which these lead of the necessity of goals and patterns for counteracting the unassailable immensity of exhaustive, random searches. Particularly important to me was the way I found this principle illustrated in a work of fiction, in the young boy’s discussion with the old hermit in Alexandros Papadiamantis’s short story “The demons of the ravine” (“*Ta daimonia sto rema*”). Here the point is beautifully made, that those who do not know where they are going are bound to get lost.

D. Last but not least, my own work as a novelist. Time and again, I – like the boy in the tale of Papadiamantis – have seen my heroes and heroines get lost in the course of the writing, as a result of unclear, unfocused, undefined or unknown goals, both long- and short-term. And the clarification of these has as a rule always been crucial in the process of editing, i.e. in the agonized efforts to transform mere description of action into story.

Two brief clarifications on the subject of intentions and goals are necessary:

The first is that a character’s goals may be partly or wholly unknown (unconscious) to him or her or even, in some cases, *not really his or hers*. This seeming oxymoron is explained by the fact that the only real person existing in a story’s world is the author[14]. Not unlike Dwayne Hoover in Kurt Vonnegut’s *Breakfast of Champions*, an author thinks that he or she is the

[14] Though, unless otherwise specified, my references are to narratives in general, regardless of the medium, I use words like ‘write’, ‘read’, ‘author’, etc. to avoid bland generic words like ‘producer’ or ‘consumer’ (of stories).

only being in his or her world endowed with free will – but unlike Vonnegut's hero, he or she is right. Despite writers' comments about "the characters taking off with the story" or "doing their own thing" or whatever – which does of course tell us interesting things about the creative process – characters in any story are really the author's avatars, and so in many cases where a goal is not a character's, it is *the author's for the character*. (Thus, for example, it is not Romeo's intention, when going to the Capulets' dance, to meet Juliet – but it is Shakespeare's.)

The second point is equally crucial: the language of goals should *not* be thought of as restricted to the higher strata of the analysis of action, nor solely to the more profound or exalted dimensions of a character's inner life. To have needs, intentions and wishes, and to be motivated to set and achieve goals via plans characterizes most of human behavior, down to the most elementary action – in fact, it characterizes action, period. Thus, though an action may often be without a higher motive – at least not a visible one –, and certain things can happen to a character without having been intended by him or her, even the simplest, most mundane actions are as a rule goal-oriented. Thus, for example, the sequence of actions involved in getting, filling and drinking a glass of water, are serially structured in time, the result of the wish to execute a plan, to fulfill a goal, that caters to satisfying the need to counteract thirst. Jay Gatsby's grander goal in the novel which bears his name may be the love of Daisy Buchanan, but in its progress he, like any character of any story, has and achieves – or fails to achieve – many minor goals and plans, as for example planning and going through the plan of taking Nick Carraway, the narrator, into town for lunch, at the beginning of Chapter Four. (Of course, this sub-goal fits in well with his main goal – but others may not, they may be independent or even antagonistic to it.)

Action, like walking, evolves step by step by step. But as – unless we are following a particularly curved path – we mostly walk in straight lines, so action consists of meaningful units of more than one step at a time. These are, as a rule, at least to some extent planned.

3.2 Action and choice

At the root of narrative intelligence is the description of action or, more abstractly, of phenomena that develop in time, of *becoming* rather than *being*. And, as *categorization* (i.e. being able to perceive, understand and use a concept such as "cat" or "table" or "home") is the essential cognitive tool for dealing with the multiplicities of *space*, so is the *action sentence* (e.g. "the cat chases the mouse" or "Alexander conquered Asia") for capturing the ceaseless flow of *time*. Put a few action sentences together, with some extra conditions, such as focus, some spatiotemporal continuity and some causality (an extension of Aristotle's *unities*), and you have a story. But the basic element is the action sentence.

It is for this reason that comic books are a particularly good form through which to study some of the essentials of storytelling: each panel depicts a basic action, i.e. *illustrates an action sentence*[15]. In comics, the *atomic* nature of basic actions in storytelling is thus brought home in a much stronger way than in natural language, using which we may be tempted – having suffered at least a dozen years of a school education – to further break down sentences in their grammatical parts, which however have no direct analogue in time[16]. But storytelling evolves in time.

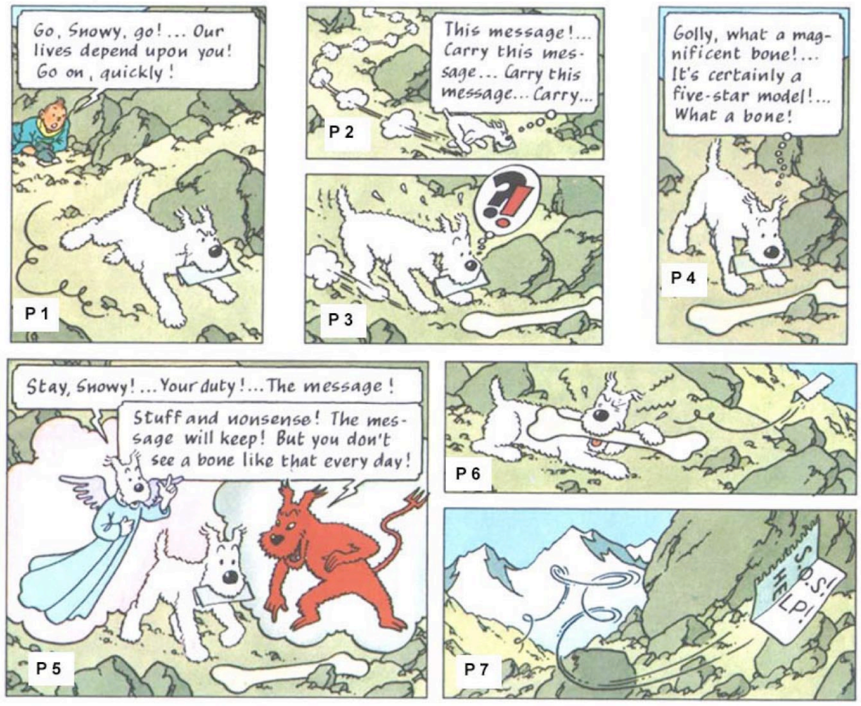
Let us look at a half-page scene from the comic book *Tintin in Tibet*, by the Belgian master Hergé. As in most comic books, here too the panels are meant to be read in linear order, one after the other – as sentences are read in a written story –, a process so essential to the form that the great comics artist and theorist Will Eisner defines comics as "sequential art". (This sequence is here indicated, rather pedantically, by the ugly panel numbers, "P1" to "P7" – these are obviously my own.) [17]

---

[15] On pages 75-77 of Scott McCloud's masterpiece, *Understanding Comics*, there is a sort of informal statistic whereby the transition from panel to panel is found to be of the 'action-to-action' variety in a percentage of 70% in almost all of the Western-style comics studied – the percentages are slightly different for the Japanese Manga, for stylistic reasons mostly –, with about 25% being 'subject-to-subject' and the remaining 5% scene to scene.

[16] Thus, though in the conjunction of two action sentences, as for example "when I went to the house I saw Marina" we can temporally separate "going to the house" and "seeing Marina", we cannot give a direct temporal analogue to the words "I", "house", "the" – or even to "went" or "saw" without their context. This further analysis we can call, as far as narrative is concerned, *sub-atomic*.

[17] The reason the fifth panel (P5) contains more information than a mere action sentence has to do with the fact that it contains dialogue – and thus is really the conjunction of two action sentences. In fact, we can say that P5 is in some sense *molecular*, a first-level combination of atomic action sentences.



I have chosen this particular sequence of panels because it shows in a very evident way something that, though it happens almost all of the time in stories, is usually not paid attention, unless you are looking for it: *a character making a choice.* And many, if not actually all, important actions in stories – often called, most interestingly, *turning points* –, as a rule involve a character facing a dilemma and making a choice which seriously affects the story's development.

The plight of Snowy in this scene is typical of a basic situation in works of the human imagination. There are countless examples: Hercules meeting in his youth Virtue and Pleasure and having to choose from the two diametrically opposed life scenarios they offer him; the warrior Arjuna deciding whether to fight, or not to fight, Dhritarastra's sons, in the most famous scene of the *Mahabharata*; Macbeth deciding whether to kill or not to King Duncan, to become king himself... But not all scenes of choice-making are so well known – or, for that matter, so dramatic. In fact, though the situation of "now the character has to make a choice" is not so blatantly obvious in them as it is for Snowy in our example, there's hardly *any* scene in *any* story which does not include characters making choices as they confront evolving reality, choosing something rather than something else either a course of action, actually realized or encoded in speech, or – when the possibility of action is not immediately available –, an emotion with which to react to events.

Stories code information and information depends on choice. (What would be the information given by the flip of a coin if we knew both sides to be heads?) And it is through choices that this information interacts with story-characters' lives, with *x* happening whereas *y* or *z* could have happened.

Though many of the characters' choices may be all but indiscernible in their minuteness and/or triviality, they are nevertheless there, because all action proceeds through choice. And when this choice does not just involve the action of the very next moment, but also one after that, we can already speak of a *plan*.

This is one of the basic reasons why stories are knowledge-tools, why they are such good devices with which to effectively transmit, store and handle human-related information: apart from being "moving", "scary", "funny" – it is qualities such as these which we usually ascribe to them in quotidian discourse --, they are excellent simulations of human beings copiously making their way inside a challenging, often dangerous, complex environment, almost constantly having to make choices. One of the more important reasons why stories matter to us is that, through the choices their characters face and make, they help us, in all sorts of ways, with our own. The action(-s) they describe may be more or less "important" (*spoudaia*), as Aristotle demands of the action of tragedy, but it is not just the momentous events which involve choice. You do not have to be Hamlet, or Snowy, to face dilemmas. Aeschylus in the *Oresteia* calls it a divine decree, i.e. a universal law: *pathein ton erxanta* -- "he who acts must suffer". But all life is action and thus all life is suffering, at least in the sense of the constant upsetting of the stasis of some theoretical, actionless passivity. And, actually, even passivity is a form of action – a mathematician might use the beautiful phrase "trivial case of action" in this case – involving as it does the choice *not* to act, and thus also leading to its own version of suffering. In fact, we could say that *pathein ton biosanta*: "he who lives must suffer"[18].

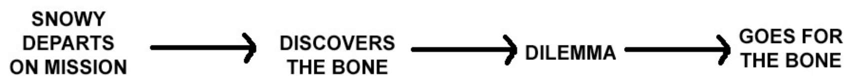
3.3 Are stories really linear?

In order to create a language in which action and choice are central, I want to start from the basic action descriptions, which I shall call *photos*. These we can also understand as action sentences or panels in a comic book. Photos, in this sense, are the atoms of stories, their temporally irreducible tiniest bits. But they do not need to describe minute actions: as an action sentence or comic panel can depict a spatiotemporally tiny event ("Bertie raised his eyebrows"), but also a more complex and longer one ("the people demonstrated in the central square") or even one spreading over a period of

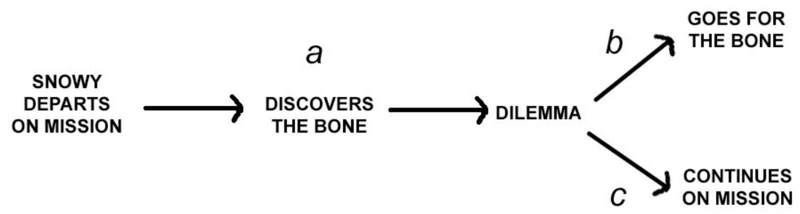
[18] It's important to stress here that these choices don't always have to be incarnated by human, or even anthropomorphic, agents: the tempest in the *Odyssey* incarnates a choice of Poseidon, while the one in the famous play of that name it's the work of a magician, Prospero. Yet the storm in *King Lear* does not appear to be the direct cause of a will – unless it be Will Shakespeare. In fact, if we wanted to create a unique, central choice-making subject in a story, we could do no better than giving this role to the storyteller.

many years and huge geographical expanses ("Alexander conquered Asia" – for this the comic artist depicting it in a single panel would need to employ iconic/symbolic visual language or even possibly also a map), so photos are basic units of action as these units are conceived in the context of a particular telling of a story, irrespective of their real-word (or *representational*) spatiotemporal extension.

Stories are written and read bit by bit, step by step, in a linear fashion, photo by photo by photo. Their arrangement can be set out in a line. Here is the main action for the Snowy vignette:



But if we look at the diagram of the action situated in the context of what *might* have happened, the unrealized alternative of the story indicated in panel 5 (P5), we depart from linearity, if ever so slightly: to describe this situation we need a second dimension:



Calling the direction indicated by the first two arrows *a*, the upper arrow after the dilemma *b*, and the lower *c*, the linear action of what actually happens is indicated by *ab*. But it is only when we see this placed in the Y-shaped diagram – that it should be placed there is indicated by P5 – and realize that it *could* have been *ac* instead of *ab*, that the story becomes interesting[19].

[19] Instances of such options and events in stories leading to multiple significant choices are wonderfully and profoundly discussed by Gary Saul Morson in *Narrative and Freedom* and Michael André Bernstein in *Foregone Conclusions*, books which their authors tell us were conceived, and develop as, twins. They deal with parallel, unrealized choice as an important dimension of some stories, giving it the interesting name of *sideshadowing*. (The first book deals mostly with the great 19th century Russian novels, and the second in Holocaust memoirs.) Though the work of Morson and Bernstein is extremely important, here we investigate a more general position, viewing every action in a story as being embedded in a larger context of choice.

If we start from any story's surface form, and instead of recording the photos of its basic actions one after the other, also draw the photos that *did not end up occurring*, those that would have happened if our characters had done otherwise – if they'd gone down, in T.S. Eliot phrase, "the paths we did not take" – we get a picture like the one in the Y-shaped diagram, describing Snowy's plight, only much more complex.

Thus, one of the basic reasons we depart from a story as a purely linear construct (i.e. as it appears in its surface form) is the fact that there are other options for the characters, which is roughly equivalent to saying that there are *other options for the action*. Marking these as available continuations for each photo, we get a non-linear structure. To make this structure become apparent, photos are not enough: we also need arrows which – in line with calling the stations of our graphs photos – we shall call *movies*; these mark the connectivity bringing forth the story's deeper infrastructure of choice, consisting both of the realized and the unrealized. It is in this that the story-as-finally-told is embedded, as a linear path.

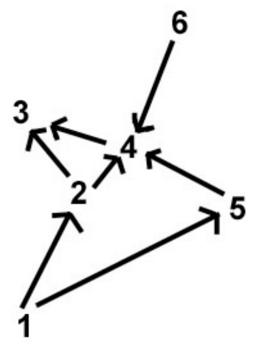
Interestingly, the concept of thinking of a story as a path inside a graph of possible events is much easier to conceive of in these days of hypermedia, computer games and suchlike. And I don't just mean because of the highly developed theoretical study of these forms of *ergodic literature*, as Espen J. Aarsperth has aptly called it[20], but for the way the *user* of these works – significantly, we don't speak of *reader* or *viewer* here – is put in a situation similar to the one in which he or she finds himself or herself in daily life, only now in worlds that are more fanciful: the situation of a choice-maker. The creator of the game gives the user the possibility of choosing – not at the same time of course! – practically *all* of the alternatives that a traditional story's hero is usually given at any point, but of which s/he can choose only one. Each playing of the game is a linear narrative that the user creates,

[20] From the Greek *ergon* (work) and *hodos* (road, way). (Incidentally, the word is not related to the same term in mathematics, statistics and physics.) Aarspeth defines, ergodic literature, more precisely, as that in which "nontrivial effort is required to allow the reader to traverse the text." And continues: "If ergodic literature is to make sense as a concept, there must also be nonergodic literature, where the effort to traverse the text is trivial, with no extranoematic responsibilities placed on the reader except (for example) eye movement and the periodic or arbitrary turning of pages." (In his *Cybertext: Perspectives on Electronic Literature*, Baltimore, Maryland: Johns Hopkins University Press, 1997). But, I note, though this progress can be trivial for the reader, it is rarely, if ever, trivial for the characters – or the author, for that matter.

based on the interactive algorithm for the navigation of the non-linear, underlying story world, created by the web-designer.

3.4 Stories as webs of photos

Let us assume that we are looking at a very minimal story whose whole world consists of six photos, i.e. including both the events that actually occurred in it *and* those which represent the various interesting eventualities which did not happen. Let us randomly assign to these the numbers 1, 2, 3, 4, 5, 6, regardless of their actual place in the events. Now, unless this is a totally deterministic collection of photos, where the first photo is precisely determined as such, and each one admits exactly one as continuation – clearly, such a scenario is totally unrealistic – linearity is not enough. To make the various photos really belong to the story world, we would need to transform our set of dots into something like the following:



Assuming just two formal rules in this symbolism, i.e. that the photos of a story can only connect in the direction that arrows departing from them are pointing, and that paths have to stop if they reach a cul-de-sac, and not before, here are some possible stories:

1 – 2 – 4 – 3

6 – 4 – 3

1 – 5 – 4 – 3

In fact, the graph is enough to tell us what *all* the possible stories containing these photos are, i.e. to generate all of them. So: photos *and* movies, define stories.

We must add here that this small story graph, like any story-world in reality, does not need to delimit a reader's imagination, or be the same for one reader and another. Thus, assuming photo 3, which is the only available ending in this graph, to be sad, a reader can easily imagine a happy one (call it 7) and have movies leading to 7, instead of 3. That 7 will not eventually occur in the story, is something that will affect this particular reader's perception of the story, and it is only in the context of it existence that his or her reading makes sense.

This is one of the reasons why it is impossible to usefully restrict, except in very artificial situations, the graph of a story to something small and simple. To return to T.S. Eliot and "Burnt Norton": "What might have been is an abstraction / Remaining a perpetual possibility / Only in the world of speculation." But it precisely in this world of speculation that stories live, and it in this we must enter if we are to understand how they work. For like icebergs, stories show only their tip, the linear structure, a surface path. But the world of speculation surrounding the photos of even the simplest story is immense. Rather than attempt to reconstruct it in its totality, a futile venture, we can just think of it as embedded in the humongous graph of all potential events, i.e. to see it as a *subgraph* of what I'll call E-space (E from *event*), which we can think of as the womb of all stories. Like Georg Cantor's "set of all sets", E-space is both inconceivable in its entirety, and most probably prone to all the logical diseases that destroyed Cantor's behemoth. But this is immaterial, as we will not have to be formally introduced to it, or refer to it except as a source of infinite potential, a context in which even the biggest story graph is but a mere speck.

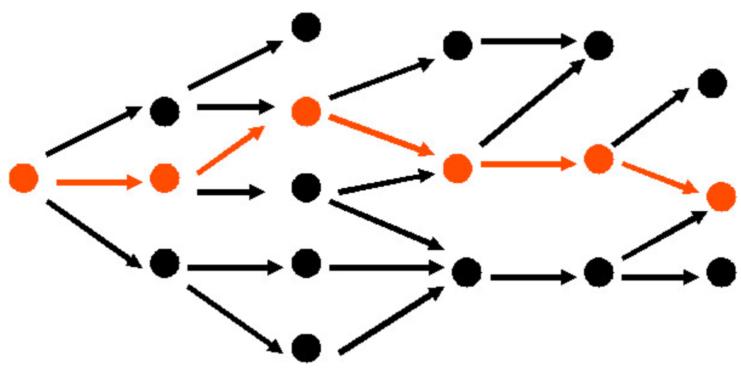
Speaking of the story graph of any story as a subgraph of E-space helps us to always think of it as located in a space of options and connections that is as complete as possible, and then reduce it to proportions that suit us better, either by the limitations of our knowledge – but knowledge can grow –, or things like the prejudices of the particular *Weltanschauung* in which the author is operating, or the conventions of a genre. (Thus, if one is writing a classical whodunit, the option of the killer not being discovered is not really available, to either writer or reader – but this lack of resolution could occur in a modernist work playing with the mystery genre, such as Paul Auster's *City of Glass*; or, in a naturalistic novel, we cannot consider the option of an assassinated hero being physically resurrected, unless it be in another character's dream or delirium.)

It is useful to recall here the distinction highlighted by the Russian Formalists, who distinguished between the story *as it happened* -- perhaps it would be more correct to say *as it happened or as if it would have happened, if it was true* –, in other words in the natural temporal order of events and in their most complete rendering, as the *fabula;* and the particular choice of events and their temporal (re-)arrangement in a telling of it – using flashbacks, for example –, as a *syuzhet*.

In accordance to this terminology, we can call the subgraph of E-space which very comfortably contains the fabula, and all the alternatives to its photos, including the connecting movies, as the *Ur-fabula*, the womb from which the fabula eventually emerges. Of course, any particular story's Ur-fabula lives in a tiny neighbourhood of E-space – but it's important for us to remember the greatest universe in which it belongs, for the eventuality that we want to transcend it. This would happen in, say, mixing different stories, like the story of the Arc of the Covenant and Indiana Jones in *Raiders of the Lost Arc*, or Sherlock Holmes and Dracula, et al., in Alan Moore's *League of Extraordinary Gentlemen*; or in the transposition and variations of an old story in a new context, as for example Sophocles' *Antigone* becoming Anouilh's, or Shakespeare's *Hamlet* becoming Disney's *The Lion King*.

3.5 What's in a movie: glue, time and causality

We started to develop the graph language of photos-and-movies also in order to be able to talk about quests and goals, concepts which become more meaningful if we look at the non-linear structure behind a story's surface linear path. Up to this point, we saw this non-linearity as being due to the fact that on certain occasions we have more than one movie (arrow) coming out of a photo (dot) as in our Y-graph of the Snowy dilemma. In this sense, a story's Ur-fabula might look more or less like what graph theorists call a *tree*, i.e. a graph beginning at one photo, the *root* (here this is the beginning of the story) and constantly expanding at every photo, to end in multitudinous branching. One of the photos of this last level of branching would be the last photo of the story, and a linear path inside this huge tree leading linearly from beginning to end would represent the story-as-published – in this example not unlike the red path:



Actually, the graph in this example deviates from the graph-theoretic definition of a tree, in that we have cases of more than one movie ending at one and the same photo. (More on this below.) But our final surface story, our path, is totally linear, photo after photo after photo, all connected in the original graph.

Now, though thinking of story worlds as photos-and-movies nestled inside some Ur-fabula allows us to explore the world of choice, and thus also intention and goal (which are about limiting choice) in a more concrete way, this non-linear language also leads to a particularly interesting way of looking at causality.

To see this, we first have to ask ourselves about the nature of the arrows in story graphs, the connections of photos that we called *movies*, beyond and before any kind of formal definition. It is clear that *photos* have a direct cognitive analogue: it is the *action sentence*, this minimal arresting of time, the temporal equivalent to the cognitive function of dealing with spatial multiplicity and variability via categorization.

But what is the cognitive analogue of the movies?

To begin to answer this question, let's first try to see if movies have any kind of local habitation, if they appear, wholly or partly, in the linear, surface form of a story – as photos so evidently do. Photos are directly represented in the surface form, either in sentences (if the surface form is textual) or in panels if it's, say, a comic book. If, for example, we take the photo that is verbally depicted as "the cat chases the mouse", we could translate this into comic book language as:

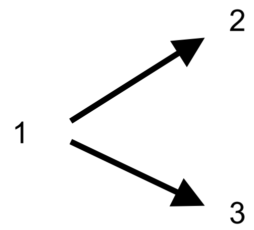
If we view it as the beginning of a mini story – and not opting for originality -- there are the two basic options for its continuation, lodged in a minimal story graph with three photos:

(1) "The cat chases the mouse."

(2) "The cat catches the mouse."

(3) "The mouse escapes the cat."

Let’s take the simplest case, where photos (2) and (3) are mutually exclusive. (Of course, more complex narratives can be constructed by their composition, such as “the cat catches the mouse and then the mouse escapes it” – or Tom and Jerry would be out of business.) The way to represent this situation in photos-and-movies language is:



If we choose a tale with a happy ending (for mouse lovers), the story is 1-3, the lower path. But as this lives in the V-shaped space of the graph, 1-2 could theoretically be another possibility, and it is this possibility that makes 1-3 a happy tale.

Now, let’s take the path 1-3 and see how it translates into story, always sticking to the comics medium:

Here it is then, one photo and then another, a basic little story. But where’s the movie?

If we imagine a similar action in the highly mimetic narrative form of the cinema, we might be excused for thinking that a film of this action would actually describe also *how the action goes from 1 to 3*, and thus contain our movie. But, if you think about it, even in a film we don’t need anything but a prolongation of 1, as the sentence “the cat chases the mouse” can cover also a longer, and more complex chase. In fact, the last frame of sequence of shots representing photo 1, in the film, would be followed by the first frame of the first shot of the sequence of shots representing 3. So, again: *where is the movie?*

To find traces of it, we have to go back to verbal language, the original tool for storytelling. Let's look at a very bland, unimaginative prose transcription of the Hergé vignette we saw earlier, broken down into action sentences (the corresponding panel numbers in parentheses) [21]:

(a)	"Snowy departs to deliver the SOS message."	(P1, P2)
(b)	"<u>On the way</u>, he finds a bone."	(P3)
(c)	"He is tempted"	(P4, P5)
(d)	"<u>And after</u> some internal turmoil"	(P5)
(e)	He decides <u>finally</u> to go for the bone."	(P6)
(f)	"<u>And so</u> Tintin's note is blown away."	(P7)

Here, at last, we can see traces of movies: they appear in the underlined words, conjunctions, prepositions, adverbs, adverbial or prepositional phrases. In their appearance in these sentences, however fragmentary, we can understand one of their functions that has a direct cognitive analogue: movies are *connectors*, a special form of glue, sticking photos together. What in the comic book form is invisible, falling as it does between the panels, in the emptiness of what comic book artists call the *gutter*, leaves a visible trace in language[22]. (Of course the Hergé sequence of panels could be transcribed in a way hiding the movies. As an example, see Jacques Prévert's "Déjeuner du matin", which, for stylistic reasons, contains almost no trace of movies, giving the concatenation of action sentences of the poem a syncopated, harsh feel.)

---

[21] That *d* does not contain a verb does not make it any less of a photo. It is one, as can be seen in its transcription "Snowy has some internal turmoil". The fact that the correspondence of panels to prose description is not one-to-one has to do both with the capacity of language to compress or expand information at will, as well as with the particular prose version – it would be simple to record the story in words in a way in which we had a precise correspondence, even if it turned out to be more pedantic.

[22] In the spirit of this last remark, it is particularly interesting to refer to Scott McCloud's theoretical discussion of the importance of the gutter in sequential art, in pages 64-69 of *Understanding Comics*. McCloud insists that this jump from panel to panel, over the void of the gutter, is what constitutes the very essence of the medium, giving as it does to the reader the locus for his or her own creative involvement in the narrative, by filling in, consciously or unconsciously, the action that connects the two panels. Whether this is as fundamental as McCloud says for the form in general, is not the point here. What's particularly important for our argument is the light it sheds on the nature of the movie: the fact that it indicates a passage does not rule out the fact that it can contain – which is perhaps equivalent to saying that it can *mask* – a lot of action which is implied, or even purposely hidden, in the surface form of the story.

Yet, a movie's role as glue in the ethereal world of a story's Ur-fabula masks two very different underlying dimensions, both of which are crucial to the construction of a story out of a sequence of photos, and which are sometimes – but not always – traceable in its linguistic form. These are:

(1) The temporal, or sometimes spatiotemporal, continuity of the action ("this happened after that")

(2) Causality ("this was caused by that")

These two functions are not mutually exclusive: every movie can contain both, in varying doses[23]. But let us, for reasons of clarity, look first at the two dimensions separately, assuming for now that every movie is either of a temporal (or spatiotemporal) type, in which case we can call it *type-1*, or a *t*-movie, or of the causal type *c*, which we'll brand *type-2*, or a *c*-movie. We can write these relations symbolically as E_1tE_2 ("E_1 *t*-movie E_2" or, more informally, "E_2 after E_1") or E_1cE_2 ("E_1 *c*-movie E_2" or, "E_1 causes E_2"). Or, in diagram language:

$$E_1 \xrightarrow{t} E_2$$

$$E_1 \xrightarrow{c} E_2$$

Thus, in the example of the vignette, Snowy goes on his mission because Tintin told him to (thus: P1*c*P2), finds the bone because it is on his way (P2*t*P3), but decides to eat it because he makes a decision to (P5*c*P6), and the paper is blown away because of the wind (P6*t*P7).

To understand why the last of these is a *t*-movie and not a *c*-movie (at least in the binary case) we must remember that what makes a sequence of photos – rather than a random concatenation of actions – a *story*, is that it is *focused* on action and/or character and/or subject, a focusing which also creates a criterion of relevance for the photos appearing in it. Thus, we wrote P6*t*P7 for the last movie because the reason the note was not delivered (P6 shows this clearly) was *not* the wind, but Snowy's abandoning the note for the bone; but had the wind been the main reason for Snowy's failure to fulfil his mission, it would have operated causally in the story. In both cases, of course, P6*t*P7 or P6*c*P7, the wind is the wind, and P7 could have the same content. But it is not the action (photo), nor the fact of the connection of two photos via a movie, which determines causality but – as a rule – the degree of

---

[23] I mean here more than the obvious fact that things that are caused by something happen after it – but there will be more on this, important matter, in what follows.

relatedness to the story's focus, be it character- and/or action-centered[24]. Thus in the *Odyssey*, say, the storm following Odysseus' comrades' opening the bag of Aeolus *is* causal, because it results from their irresponsibility[25].

* * * * *

The narrative mode is our basic tool for understanding time-dependent existence and, above the atomic level of the action sentence, time in narrative advances through movies. Thus, to understand their way of operation is to understand the quintessence of narrative, at least at the molecular level of analysis, i.e. of how t photos combine to form larger wholes, at the first step of complexity.

Let us first look at their temporal function, i.e. the *t*-movies. These are much simpler in their action. One of the reasons for – or possibly a symptom of – this is that they are *strongly linear*, a property which is pretty much synonymous to the fact that they are totally transitive; in other words:

$$(E_1 \xrightarrow{t} E_2) \quad \text{AND} \quad (E_2 \xrightarrow{t} E_3) \quad \text{implies} \quad (E_1 \xrightarrow{t} E_3)$$

In the cognitive sense, this is obvious: unless we are in a science fiction novel set in a universe with a totally different notion of time, we can always say that if "E_2 happens after E_1" and "E_3 happens after E_2", then "E_3 happens after E_1."

The fact that *t*-movies are very simple in their operation, can also be seen in the "and then" function of basic storytelling, prevalent in very young children. (The skill of serial, *t*-ordering of photos, is developed much earlier in life than the more advanced, of *c*-ordering.) Although causal linking is also an extremely important dimension in narrative, the *t*-element is crucial in a very fundamental kind of way: in fact it is the ground on which causality is built, developing on the principle of *post hoc ergo propter hoc* ("after which, therefore because of which").

---

[24] It is for this reason, that the strong doses of coincidences and miraculous escapes in the ancient Greek novels as well as in Victorian melodrama, and many other forms of popular literature, is frowned upon, as inferior plot construction. But it should be said that in both, in the first due to the veneration at the time the Greek novels were written of the goddess *Tychē* (Fortuna in Latin), and in the second in some rather unfounded theologically sense of divine providence, these seemingly fortuitous events were attributed to a more profound, controlling power of human affairs.

[25] This is another good point to emphasize the validity of seeing Ur-fabulas of stories as subgraphs of E-space: one and the same blustery photo (like P7) can find totally different interpretations if the movies departing from it can extend far enough, into the part of E-space containing elements of an archaic cosmology.

But this is not the sole function of *t*-movies. Apart from their important role of ordering-via-gluing, *t*-movies are the most essential tool of *realism*, i.e. the way in which narrative convinces as an imitation of an actual world, lying out there. One of the most important characteristics of lived experience is its continuity – at least at the small temporal scale – the fact that we mostly perceive action as smooth and flowing. And it is this element that the narrative mode imitates by gluing two or more events by virtue of their temporal consecutiveness and – if there is not a *c*-factor also involved – nothing else. The sentence “I woke up, then dressed and then had breakfast” is not out of Proust (though it could be!) but it is amazing nevertheless, in that it connects three different actions though there is no deeper relationship linking them, other than the fact that they happened the one after the other.

* * * * *

The other function of movies, i.e. *causality*, is especially important to our analysis, for it is this that mostly supports the all-important dimension of wishes, intentions, motivations, goals and plans: most of the causal connections in stories evolve out of the wish of some character, the wish transforming into intention, intention into goal, goal into plan to fulfill it. It is *c*-movies which form the mostly invisible spine of the macro-structure of a narrative, the connecting principle that can take us all the way from the atomic level, to the molecular, to the cellular, eventually leading to the level of the whole story, as an organism. Macbeth kills Duncan because he – and the wife, of course – wants to be king, the inveterate romantic, Jay Gatsby, moves to West Egg and starts giving all the wild parties because he wants to get close to Daisy, Sylvester chases Tweety because he wants lunch …

We do not *just* perform acts, we perform them as parts of larger acts, that are lodged in time-dependent scenarios. Even routine is planned – in fact especially so: unless we are particularly neurotic that way – which is another story –, we do not start to ruminate on whether we should move house every single morning, as we wake up, and only decide against it after a call to real estate agents, and the receiving of reports that the market is high. This is because we have a *plan* to stay in the same place for a while, though this is almost always taken for granted and thus not conceived of as such. (Again the mathematically inspired expression “trivial plan” comes to mind.)

Of course life is full of surprises – but they are surprises precisely because they do not conform to the plans. *Of course* there is place for the random and the contingent, and cases of brief exemptions from some of the demands of time – but they are that: exceptions. And if strong doses of randomness and/or contingency in the real world around us on occasion make us think that chance predominates, let us remember that *life is not* story. In fact, the reason that the contingent and the random play a smaller part in stories than in life – except in some cases of experimental fiction – is a basic part of the their *raison d' être*: connecting photos causally in movies is not just an attempt to display the deeper structure of the world but also an *injection into it of extra doses of causality*. This injection is performed by the storyteller either in the name of a deeper understanding of the world ("showing how things really are") or of an artificial construction, reading more causality than there actually is in things in order to give extra meaning to an otherwise absurd world, and/or extra zest to a narrative's appeal. The line between the two, of course, is extremely difficult to draw, a thing which storytellers are notorious for exploiting.

Causality gives meaning to action, and it is the existence of goals that forms its organizing principle. Take away goals and plans, and any causality left in a story will be extremely short range, and thus reduced to the random games of instincts and natural phenomena, the people in it moving like Marcus Aurelius's puppets jerking on their strings (*sigillaria nevrospastoumena*). And life will be a tale told by an idiot, signifying nothing.

Of course, all *c*-movies are also *t*-movies, as this is the way our world operates: the cause temporally precedes the effect. *But not all t-movies are c-movies*. In fact, "*post hoc ergo propter hoc*" is most often used to denote a well-known fallacy, i.e. of *always* mistaking *post* for *propter*, a fallacy which we meet as a form of magical thinking dominant in small children ("the sun rises because the cock crows" or "there was a thunderstorm after I had bad thoughts about mommy, therefore it was my bad thoughts which caused the thunderstorm"). Yet having a fallacy named after it should not blind us to the fact that in many cases in a narrative – there is not an equally high correlation in real life – *after* does indeed mean *because of*: in the narrative mode of thinking, spatiotemporal proximity is the strongest and most basic, though not the only, indicator of causality[26]. For example, in the following two cases we don't need a "therefore" or other connective – though they could well be used – to denote the fact that there is a causal connection between the first photo and the next:

Jim pushed John. John fell down.

Nikos called Alexandra a loser. Alexandra cried.

Causality connects photos in a way which goes beyond mere temporal concatenation, to denote a more profound connection, a sense in which the later ones are a result of the earlier. But the causal connection itself, the *c*-movie, can often be invisible: "the movie lies in the gutter", we'd say borrowing

[26] I had not realized how fundamental this mechanism of temporal proximity as the basic evidence of causation was until I read in Malinowsky that the causal connection between coitus and pregnancy was not at all obvious to members of the pre-literate society of the Trobriand Islands.

comics parlance. Thus, though the two action sentences in each case are merely concatenated in the previous two examples, we infer its existence both by virtue of content, as well as our tendency to read *post hoc* as *propter hoc*, educated as it is by knowledge of the ways of the world.

Yet, the special nature of *c*-movies, and their difference from *t*-movies, becomes really apparent the moment we wonder about their transitivity. For though causality appears to be transitive in the short term, and thus combining "John fell because Jim pushed him" and "John hurt his head because he fell" makes a reasonable case for "John hurt his head because Jim pushed him", if we take longer strings of it, the transitivity breaks down. In fact we can say, very informally, that *causality is not very transitive.*

Look at this statement: "Dresden was destroyed because Gavrilo Princip suffered from tuberculosis." Even to those who remember who Gavrilo Princip was, this statement sounds absurd. Yet, see how it can result from a sequence of *c* connections – please notice that there is a clear causal link from every photo to the next.

(1) "Gavrilo Princip (a Bosnian) was suffering from advanced tuberculosis."

(2) "Gavrilo Princip had little time left to live."

(3) "The head of Serbian Intelligence chose Gavrilo Princip as a suicide attacker of Archduke Franz Ferdinand."

(4) "On June 28, 1914, Gavrilo Princip assassinated the Archduke."

(5) "WW I began."

(6) "In the 1920's there was economic decline, poverty and social unrest in Germany."

(7) "Many people in Germany were eager for a 'savior', who would restore their nation's honor, power and wealth."

(8) "Adolf Hitler and the ultra-nationalist ideas of the Nazis had great popular appeal."

(9) "The Nazi party became a basic player in the German political scene."

(10) "In 1932, President Hindenburg appointed Hitler Chancellor."

(11) "Hitler became a dictator and assumed absolute power."

(12) "In 1939, Germany attacked Poland."

(13) "WW 2 began."

(14) "The Luftwaffe bombed London."

(15) "The Allies bombed German cities."

(16) "Dresden was destroyed."

So, if every *n+1* in this sequence occurs “because of” *n*, this sequence, why is it totally absurd to say that (16) occurs “because of” (1)? There are basically two related reasons, which we can express in graph language in the following way: a) *c*-movies are not, like *t*-movies, zero-one phenomena, and b) more than one *c*-movie can arrive at a photo.

In fact, given any two photos *a* and *b* in a story, there are only three possibilities regarding their temporal ordering: that *a* is after *b*, that *b* is after *a*, or that they are exactly simultaneous – but as the third case is extremely rare, and when it occurs in a story it as a rule has a some special significance, we can focus on the first two and say for simplicity that *t*-movies are (in almost all cases) indeed zero-one. But seeing how causality operates in the world, we must think of the causal factor as graded or weighted, in ways in which the temporal isn’t [27]. The underlined words in sentences such as “I was <u>partly</u> influenced by this”, “it was that <u>and other things</u> that caused Mary to act the way she it”, “for Petros, Marina’s answer was <u>the straw that broke the camel’s back</u>”, or similar, makes this clearly apparent.

So, let us give every *c*-movie *m* a *weight*, i.e. a number w(*m*) measuring the extent to which it is causal. To depict causality realistically, the weight of *m* can range in the full interval, from 0 to 1, i.e. $0 \leq w(m) \leq 1$. In the case where we have $E_1 m E_2$, with E_1 occurring at t_1 and E_2 at t_2 , and though $t_2 > t_1$, w(*m*) is 0, this will mean that – chaos theory-type talk of butterflies and storms notwithstanding – the occurrence of photo E_2 is *totally independent* of E_1: in fact, we would immediately deduce that *m* is a pure *t*-movie. At the other end of the spectrum, if w(*m*) is 1 this means that the occurrence of E_1 *absolutely determines* the occurrence of E_2, i.e. that if E_1 occurs E_2 must *necessarily* follow. But weight values between 0 and 1 have in-between degrees of influence, i.e. *E_1 happening partly determines E_2 happening.* We know that life is rich with such examples. In fact, even in the much more causality-determined world of stories, movies with *c*-value of 1 are quite rare – though of course they exist[28].

To give an example from the case of the Gavrilo Princip path: it is clear that the *c*-movie connecting (1) and (2) has much higher weight than that connecting (2) and (3): the fact that Princip was suffering from advanced tuberculosis made it very likely that he had little to live. But that this was so, does not lead with an equal degree of certainly to the fact that the head of the

---

[27] A sophisticated mathematical analysis of causation is presented by Judea Pearl in his book *Causality*. But as here we are not concerned with the scientific analysis of causality, but the *folk* level of its understanding (as in “folk psychology” or “folk physics”) which is the one dominant in human narrativity, we can use a much simpler model.

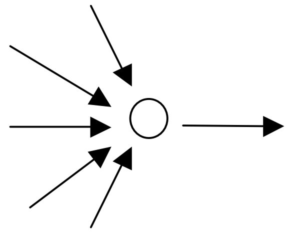
[28] As such cases of movies carry minimal amounts of information (like the flip of a coin with heads on both sides), they are less interesting. Regarding determinism and non-determinism in narrative, and its philosophical significance, read the profound analysis of Morson in *Narrative and Freedom* and Bernstein in *Foregone Conclusions*.

Serbian Intelligence chose him, especially, as the assassin: no doubt, there were many thousands of seriously tubercular people at that time in Bosnia and the weight of this causal link is inversely proportional to their number. The reason why Princip, and not another, was chosen, undoubtedly also had to do with other factors, as for example that he was young and enthusiastic, that he had a special kind of character, that he had strong nationalist feelings and, more especially, that he had happened to be – perhaps originally for *t*-reasons – close to the "Unity or Death" nationalist organization.

One of the great advantages of weighted *c*-movies, where the weight measures, in everyday terms, the extent to which one event (photo) influences the occurrence of another, is that they make it much easier to compute the causality of paths. This weight in essence being a probabilistic measure, the causality of a path of *c*-movies is the weights of the individual movies multiplied[29].

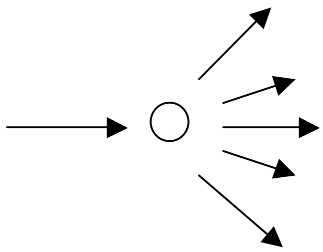
The second reason why transitivity breaks down reflects an important characteristic of the operation of *c*-movies: more than one *c*-movie can arrive at any one photo – in fact, as a rule many more do –, as in the example of Princip being chosen as an assassin for a combination of more than one reason. And it is this fact which explains why it makes sense to speak of a *c*-movie acting effectively on a photo, though it may have low weight: *because it acts in unison with others*. In fact, there is an additive law of sorts operating. The weight of *c*-movies terminating on a photo act on it according to the sum of their weights with 1 as a maximum (i.e. in this particular arithmetic, values over 1 are reduced to 1.) Thus, if every incoming movie in the following diagram has c-weight 0.2, their totality will be 1: and all five movies acting together on *n* make it *certain* than it will happen.

[29] As an illustration, even assuming an 80% degree of certainty (0.8 weight) of each of the movies connecting consecutive photos in the Gavrilo Princip example – and it's definitely much much lower than this in many, if not all, the movies – the causal influence of the first to the last photo of this sequence, i.e. "Dresden was bombed because Gavrilo Princip suffered from tuberculosis", is around 2%. And if every connection has 20% influence (0.2 weight), which seems closer to reality in most cases – and is way below that in some cases, as for example the link 2-3, where it could be down to the order of 10^{-4} – we go all the way down to a degree of 10^{-12} for the influence of first to last.



This diagram is important as it shows, at the atomic level, the second reason why a story graph is not linear: the many-to-one action of some *c*-movies (essentially, the photos they come from) on one and the same photo.

The many-to-one local structure, shown in the following diagram, we saw in 2.3, *à propos* of the fox terrier's dilemma:



One of the consequences of multiple in- and out-going movies related to one and the same photo, is that its function in a story can vary immensely, depending on which of the movies related to it in E-space are chosen. And this can also be determined by which subgraph of E-space the photo is seen to belong to. Thus, a decision in a man's private life can be important if we are looking at his nuclear family's story's subgraph of E-space, extending to a rather short period, but may be less so in K's extended family's story's subgraph over a longer period, and not at all in K's city's story's subgraph. But on rare occasions, even a small personal decision can affect the immense subgraph of world history – for example, Princip may have contracted tuberculosis by going out with a tubercular girl, a small private-life decision that had a contribution to events of momentous importance. (Or for that matter, look at what happened when Paris, prince of Troy, laid eyes on Menelaus's wife.)

To close this discussion of causality in stories, let me comment on the way it reflects on the use of the everyday sense of "*x* causes *y*", pointing to two different meanings to it, leading us to distinguish between a *strong* and a *weak* sense of causality:

- The *strong sense* we reserve for "*x* causes *y*" in the sense of "*y* is a necessary consequence of *x*" (for example in "John died, because Jim shot him through the heart").
- The *weak sense* would mean "*x* caused *y*" in the sense of "y lies on a c-path of which x is an earlier photo" (as for example in the case where, say, x = (3) and y = (6), in the Gavrilo Princip causal sequence.

One may be able to distinguish between the two uses, by setting some kind of threshold value for w(*m*), above which m might imply causality in the strong sense. But this does not really reflect everyday usage of the causality-related natural expressions, such as "caused", "because of", "as a result of", etc. which can be much more ambiguous, sometimes referring to the one and sometimes to the other sense, though without a clear verbal marker. In fact, our imprecise use of words describing causal relations in everyday discourse, the frequent mixing of the two types of causality, as well as the often extreme over- or under-estimation of weights of real-world movies or paths thereof, may be at the heart of the infinite flexibility and malleability of stories as cognitive tools. That we do not understand causality very well may be, in fact, one of the main reasons why stories have had such great evolutionary success, being themselves so adaptable to varying degrees of precision.

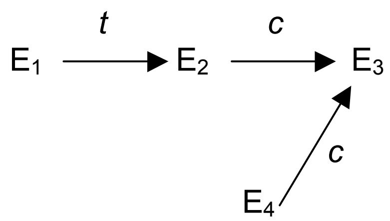
3.6 Linearization: necessary evil *and* blessing in disguise

When a sequence of photos originally appears simply as a path in the subgraph of E-space which constitutes the syuzhet of a certain story, it is very easy to translate in the linear, written form of the story. Take for example:

$$E_1 \xrightarrow{t} E_2 \xrightarrow{t} E_3 \xrightarrow{c} E_4 \xrightarrow{t} E_5$$

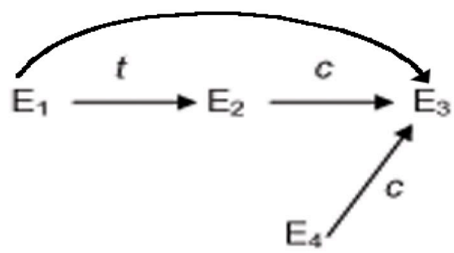
Translated into verbal natural language, the photos may be linked in the simplest way as mere concatenation "$E_1E_2E_3E_4E_5$'' (as in Prévert's poem referred to earlier) or in more elaborate ways dictated by the precise content of the photos and/or stylistic reasons, such as "E_1 , then E_2 and E_3 and so E_4 , then E_5'', i.e. forms where traces of the *t*- and *c*- movies survive in the linearization. But in any case, when we have a linear path in the subgraph, it determines a strict hierarchy (what a mathematician would call a *total ordering*) on the set of photos, and it is this which exactly determines the order in the linearized form.

But now assume that a V-shaped link occurs on a photo belonging to the syuzhet of a story, as for example in:



Even forgetting the variations on linking the photos verbally in the linear form (mere concatenation and/or use of linguistic connectors), we can see here that there are more than one ways to linearize this diagram – and linearization *is* necessary, if the graph is to be turned into verbal narrative, which is of necessity linear. One linearization might go “E_1 then E_2 and also E_4 so E_3”, but another might reverse the order of the non-linear terms, as in “E_1 then E_4 and E_2 and so E_3”. But also, one of the two E_2 and E_4 could be eliminated in the linear form – we often do this, with more than one causal factor operating in a narrative –, or it might be that both may appear, but without the c-movie linking E_4 to E_3 having left a visible trace as in “E_1 then E_2 and E_4 , E_3”. In the last example, the fact that E_4cE_3 is passed over in silence either can be due to the fact that it is implied by the rest of the linearization (syuzhet) and is thus considered obvious, or because it is not as important as E_2cE_3 or even because the storyteller purposely wants to *hide* it – as say in a whodunit, where the writer will mention a clue but wants to underemphasize its significance as such, in the case of a first person so-called “unreliable narrator” who wants to mislead the reader.

Another thing that can happen in a story, of course, is that a photo, say E_1 *can act causally on more than one photos*, so we get an enhanced form of the previous graph:



A graph of this type is no more called a tree but a *dag*, from the initials for the words “directed acyclic graphs”.

So, all of these options, and more, are in the standard bag of tricks of a storyteller. The fact that he or she can do so many different things while linearizing the inner, logical, non-linear connectivity of the photos of a story graph is one of the great beauties and a big part of the power of storytelling. Storytellers are cunning folk, and often like to play with their audience's expectations, revealing their story's truths piecemeal, when and where and if it suits their purposes, purposes that can be aesthetic and/or rhetorical[30].

The precise way of linearizing the photos of a story graph, though adhering to some basic rules, locally, has variable degrees of freedom, resulting from the nature of the process of transducing the non-linear to the linear. This, with its attendant processes of elimination of at least a few movies (either for economy or concealment) and the reduction and/or amplification of the causal factor in the process of linearization, are at the heart of narrative as cognitive mode. And many of these depend on the non-linear local structure of causality.

But, viewed from another angle, these advantages of the surface linearity of the narrative mode are but the good side of a bad coin, the positive aspects of the basically negative – at least in terms of information-processing – operation of linearization. For, as any mathematician will tell you, if a graph is above a certain level of complexity – this expression is roughly equivalent, for the purposes of this discussion, to saying that it contains over a certain number of arrows in non-linear links –, then any linear arrangement of its photos, in which each photo appears at most once, is *bound* to contain a much lower amount of information on connectivity than the graph itself. In other words, though the treasures of natural language, its abundant vocabulary and the highly sophisticated mechanisms of grammar and syntax, can provide us with ways for packing much more information into a linear ordering of photos than would be possible if we were merely concatenating, the final, linearized form of a story, unless it be unnaturally repetitive, pedantic and convoluted, *is bound to contain much less information on overall causality than the actual graph* – this information will be lost in the process of going from non-linear to linear[31]. A pertinent corollary of this is that the process of

---

[30] In probably the most notorious case of an "unreliable narrator" in literary history, Dr. Sheppard describes in the following way the scene of the murder which – as he confesses in the end – he has perpetrated: "The letters were brought in at twenty minutes to nine. It was just ten minutes to nine when I left him, the letter still unread. I hesitated with my hand on the door handle, looking back and wondering if there was anything I had left undone." The actual crime, which is probably the most significant causal link in the *Murder of Roger Ackroyd*'s Ur-fabula, is passed over in silence – in fact, it lies in the gutter between the first and the second period –, without the reader noticing anything missing. And it is interesting that the murderer/narrator later repeats the above excerpt, prefaced by the statement "I was rather pleased with myself as a writer." To say only part of the truth, in the process hiding important causal connections, is made possible by our very ambiguous ways of our very unreliable ways of linearizing complex causal graphs.

[31] An other way of saying this is that the original structure of the graph, *and especially so of its causal links,* cannot be reconstructed from the information in a sequence, unless the sequence contains (at least) multiple repetitions of some photos.

linearization cannot be done in a unique way, i.e. there are more ways to linearize a non-linear subgraph of a graph than one.

Our capacity for linearizing the non-linear worlds of story graphs and compressing the high complexity of non-linear structures into much smaller, linear ones, using language to linearly encode part of the non-linear information, and shedding the rest, as well as the reverse ability to decode such linear utterances into at least a part of their original, non-linear complexity, is probably one of the most striking features of the narrative mode.

The translation of a world of relations, both temporal and causal, whether real or imaginary, into story, involves a lot of shedding of information anyway, whether for reasons of selection, emphasis (on what is not eliminated), economy, or whatever. The action sentence “the cat chases the mouse” is extraordinarily elliptic, in its relation to any (real or imaginary) underlying Ur-fabula, and also – mostly because of this – extraordinarily to the point. But this we know already. The aspect of information loss that we are stressing here is that beyond any loss that will occur for any other reason, a huge amount of information in the story graph will be lost in the process of linearization *precisely because the original is non-linear and the final outcome (the story) is linear*.

But linearization also has its advantages. How does that old song go? “You got to win a little, lose a little…” That’s the story, *and* the glory, of narrativity.

3.7 Outlines

The existence of weak causality leads us to an observation that we can generalize to a principle applying to the story graph.

> **The Distance Principle**
>
> The farther away two photos are in time, in the story’s fabula, the lower the weight of their causal relationship (described by the weight of the shortest path of movies connecting them).

Clearly, the validity of this law is statistical, for there are rare occasions of *c*-movies, or paths of movies, of high weight, connecting events far distant in the *t*-sense. So, for example, Clytemnestra’s killing of Agamemnon in the *Agamemnon* has strong causal effect on Orestes killing Clytemnestra in the *Choephoroi*, a whole play away in the *Oresteia*. But it is no accident that these are the most important photos in the two plays. In fact, a good measure of the strength of the causal connection between two photos in a story graph is the likelihood that this connection will survive in an outline. (In our example: even the most minimalist outline of the *Oresteia* must include the *c*-connection between Agamemnon's and Clytemnestra's deaths.)

From a different perspective, the Distance Principle means that a story's causal organization is much stronger locally (at short temporal distances) than globally, global causality only operating in outlines.

Outlines are another major tool in the cognitive repertory of narrative. There is no doubt that our basic narrative skills include creating and/or understanding telescoped versions of stories and being able to relate them to their fuller forms. Of course, like many other skills, this can be refined by practice and improved by culture. Yet, the ease with which a three year-old can refer, after having heard an elaborate narrative only once, to "the story of when Theseus killed the Minotaur", or "the story about that time daddy met a real monkey" – both of these expressions are of course ultra-dense outlines – seems to point to an innate, or otherwise very basic, capacity for outlining.

Clearly, there are many levels of outlines in a story. In fact, if we start with a story of some length in its linear from, there is practically no end to the ways in which we can outline it. Also, there is no sense in which we can order these outlines totally, in a line. The length and form of an outline can and does depend on all sorts of factors, including conscious or unconscious prejudice. And for this reason, at every level of analysis (you can think of this as degree of resolution, total length of the outline, or ratio of the length of the outline to length of the story) there can be many radically different outlines[32].

In fact, if we start with an ideal as-complete-as-we-want-it fabula of a story F, the set of its outlines forms a tree structure: you can imagine F on the ground, and the tree going upwards; outlines at any level will be coarser versions of those under them, but may be much less related to the lower levels of other branches, having developed with different emphases, each following the biases of its own branch. As, especially in the oral cultures where storytelling begins, the name of the game is variation and there is really no "totally correct" version of a story, so there is no absolute sense in which a story is *a* story rather than an outline of a larger one or, for that matter, the expanded form of an ideal original. Rather, we can more profitably talk of many *levels* of a story, or the tree of outlines in which each story exists[33].

Now, assume we begin with a story S, and want to outline it. In the context of the photos-and-movies language we've created, we spoke of the

---

[32] Remember the scene in *Shakespeare in Love*, when the actor playing the Nurse is asked in the tavern what their new play is about, and he begins his response with: "Well, there's this nurse…" The play, of course, is *Romeo and Juliet*, and the actor is about to give a Nurse-centred outline.

[33] Of course, when we are referring to pubished stories, things change: thus, the original book version of *Moby Dick*, or Art Spiegelman's *Maus*, or Antonioni's *L' Avventura* have to be considered, in each case, *the* story (the syuzhet, more correctly) of these particular forms. But, again, there is no end to the ways in which we can outline them.

Ur-fabula of S, which we think of as the smallest subgraph of E-space which contains all its causal photos (even those connected with a very low weight), both realized and unrealized, as well as the photos connecting them. Calling U(S) the Ur-fabula of S, and S_n one of its outlines, we can now ask the question: what relationship does $U(S_n)$ bear to U(S)? Is it a subgraph? And if so, what kind?

We shall describe two basic ways in which the two can be related as well as a third, which results from their combination. The three together cover the range of possible answers to the question. Calling a story graph U* an *outline* of graph U(S), we can speak of three different forms: *excision*, *contraction* and *combination outline*.

Excision outline

A story graph U* is an *excision outline* of U(S) if:

1. The photos (vertices) of U* are a subset of those of U(S)
2. The movies (edges) of U* are a subset of those of U(S)
3. There is a syuzhet (path) S_x in U*, such that all photos and movies of S_x are also in the path S, i.e. the story of which U(S) is the Ur-fabula. Two consecutive ones in S_x may not be further apart in S, though their order in S must be preserved in S_x. (This means that some of the photos in S may be missing in S_x, but if *a* comes *b* in S, and both appear in S_x, then *a* will come after *b* in S_x.)

The intuitive concept behind this definition is indeed that of an *excision*, i.e. of “outlining” a graph by excising a subgraph out of it, without adding any new photos or movies, just including the “more important” photos and movies. In this case, the linear excision outine S_x of the story S will be one referring to major events of S, and their relations. Those this seems like a rather natural way to outline, the excision process can be less than completely functional, for the reason that some crucial elements of a story may not be included in it as separate photos. For example, “Odysseus spent many years on the island of Calypso” is not one, but *many* photos in the *Odyssey*. To cater to cases such as this we need:

Contraction outline

A story graph U* is a *contraction outline* of U(S) if there exists a partition of the set of photos U(S), into a non-intersecting set $PU_1(S)$, $PU_2(S)$, $PU_3(S)$ … $PU_n(S)$, which we shall call PU(S), such that:

1. The union of all $PU_i(S)$ is equal to U(S).
2. U* has n photos, each P_k of which corresponds to one $PU_k(S)$.

3. There exists a movie between photos P^*_k and P^*_m of U* if and only if a photo of P_k is connected by a movie to a photo of P_m in U.

The intuitive analogy here would be that each photo P^*_k in U* is really, in a sense, already an outline, or even a title in one action sentence, of the photos of $PU(S)_k$. Thus, the photo “Odysseus spent many years on the island of Calypso” in the contraction outline would correspond to all the photos of the subgraph of the *Odyssey* referring to Odysseus arriving and/or staying on the island of Calypso. And it would be connected to other photos, such as “Odysseus departs from the island of Calypso” because there are movies in the *Odyssey* taking us from the one to the other. Thus, the gale caused by Poseidon’s anger at the blinding of his son Polyphemus is in a *t*-relationship to the photos of Odysseus’ departure from the island of Calypso, so the larger, in terms of events, photos containing these events in a contraction outline will also be in a similar *t*-relationship in the outline.

Combination outline

A story graph U* is a *combination outline* of U(S) if it consists of a combination of excisions and contractions of U(S).

This third definition probably covers most of the outlines that we meet in actual storytelling, as different outlines will contain different degrees of resolution on any part of a story, sometimes resembling more an excision, and others more a contraction.

* * * * *

We mentioned earlier that the stronger the weight of a c-movie is in a story, the likelier it is that it will survive in an outline. In fact, this is also true of photos: the central events of stories, such as the killing of Agamemnon and the killing of Clytemnestra, would appear in any decent outline of the *Oresteia*, no matter how coarse. They are events in the actual syuzhet of Aeschylus’ play, and there is a movie connecting them there: both the photos and this movie are excised as they are, and appear in the (decent) outline. This fact, together with something we call *independence of levels*, show us why outlines are so important to the understanding of how a story is structured.

The observation on the independence of levels in any story (we are referring here to different levels in the tree of outlines of S) is really a corollary of the obvious fact that we have to *shed* information when we outline. This is perhaps much simpler to see going in the opposite process: give an outline of a story to ten different writers to develop, and they will write ten different stories. In other words: the information in an outline is not enough to expand it in a unique way.

We can express the independence of levels – again, a statistical truth -- more formally as:

The Independence of Levels Principle

Assume p and p* to be photos in the Ur-fabula of a story S, k to be the highest (coarsest) level in which a photo *p* appears in an outline branch O_n(S), and j the highest level where photo p* appears in an outline O_m (S), with j > k and possibility of m=n included. Then, the weight of the causal connection p*c*p* in the fabula, if it is > 0, is inversely proportional to some power of j – k.

We can approximate the gist of this principle in less fancy language by saying that the coarser we make an outline, the less likely is it that any one of its photos will bear a strong causal relationship to a photo in the original story. (Example: whether the character "Meyer Wolfsheim" fixed the 1919 or 1920 World Series, an event to which Jay refers to in Chapter 4 of the *Great Gatsby*, is totally independent of the way Jay will die in Chapter 8.)

At first, this seems a totally obvious, and because of this perhaps rather trivial observation: *of course* you lose information when you have to outline a ten thousand word story into a hundred words, and the first information to go in such cases is the "details". The Prince's precise instructions to the Players in Act Two of *Hamlet* are not at all relevant in their details to the way in which he will kill the King in the last scene of Act Five: the instructions could be changed, and made different, without really affecting the play too much – or they could be cut out altogether, as they are in some versions[34]. (In fact, almost every staging of the play will dispense of some parts of some scenes, or whole scenes – but we still call it *Hamlet*.) And though the story of the youth of Father Zossima, as we learn of it in Book 6 of Part Two of the *Brothers Karamazov*, does add to the sense we have of Alyosha's spirituality, and contributes to the overall effect of the book, nothing would *really* have to change in, say, the scene of Dmitri's trial in Book 12 of Part Four, if the saintly man's brother had *not* contracted consumption and gone through a religious conversion in his teen years.

The fact that we know what "really" means in the previous sentence in essence forms a basic part of my argument, as it reflects the sense in which

[34] An example is the shorter video version of Kenneth Branagh's *Hamlet*, in which both the instructions to the players, and the "What a rogue and peasant slave am I" soliloquy, which ends the scene of the arrival of the players, has been cut. It is quite clear that Branagh follows the structure of the so-called "Bad Quarto" in this, which is closer to the focused revenge structure of the lost, pre-Shakespearean version of Hamlet.

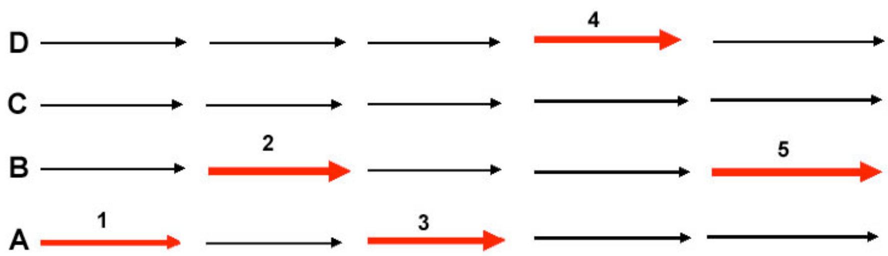
we can have a perception of a story as a whole, of overall structure, of important and less important events in it. Of course, such assessments are to a large degree subjective to each reader – but so are outlines.

But the fact that we shed information when we abstract, gives us only a part of the picture. The reason why we speak not just of *difference of levels*, which would adequately describe this shedding but of their *independence*, is that in outlining we do not just lose information but also gain. In other words, independence really means that *each level is more or less equally important in a story*, the shedding of photos as we outline both stealing information (on detail) but also giving (on structure). A basic reason for this was also mentioned in connection with the Distance Principle: strong causal narrative links survive in outlines and thus, in being outlined, a story gains in structure what it loses in detail.

An important, functional corollary of the independence of levels is that *we understand a story at many levels at once*. In fact, this is a central aspect of the power of narrative intelligence. Even on the first hearing of the an orally presented tale, a listener will encode it at many different levels. Thus, unless we suffer from some or the strange neurological syndromes afflicting certain *idiots savants*, there is no chance that we can repeat a rhapsody of the *Odyssey* word for word after hearing it recited once – or even a dozen times. But, if we are attentive, even after the first recitation we will remember pretty well the general structure, which is a very coarse outline; we will also remember in greater detail, i.e. at some finer level of outlining, some scenes which are particularly important and/or have caught our attention; and we may also remember particular words or even phrases in their entirety, reproducing on occasion some parts of the heard version more or less in the full detail of the syuzhet-as-heard. Thus, outline levels are not just convenient tools for making a story shorter, once heard – they are an essential way to process it.

What this also means is that photos from different levels of a story can be combined, to form beings which are story-like all over, without any sense of a discontinuity, as we move up or down levels of analysis. In fact, assuming some ideal fabula of a story, where all of its basic action is narrated somewhat uniformly, in the finest possible resolution (with as much detail as possible), an actual linear form of a story usually proceeds by covering different events at different levels of analysis, i.e. concatenating material that is now very detailed, now much coarser, and so on, depending on inherent interest, relevance to overall structure, the demands of meaning and focus, the likes and dislikes of a particular audience, and so on.

In other words, to give a diagrammatic illustration, if we think of a fabula as being outlined at, say, four different levels of analysis (these can be many more, of course), which we call A, B, C, D, a linear story of this can proceed in time now moving at this level, now at that, up, down, up, further down, high up, down, etc. with no apparent sense of discontinuity, as in this diagram, where we have the four levels in the vertical axis, time and action advancing in the horizontal, and the actual syuzhet is constituted by the red arrows, viewed as the sequence 1-2-3-4-5.



The independence of levels gives us extreme flexibility in analyzing, transforming and recombining them, giving us the opportunity to work at some levels, without affecting, or being affected – again, this is a statistical, not an absolute observation – by the lower (finer) or higher (coarser) ones.

That this is at all possible has to do with the underlying structural self-similarity of all levels of stories and story graphs: both in their linear and non-linear, underlying structure, stories look pretty much the same at all levels of analysis. Thus, the graph of a very rough outline of the *Oresteia*, perhaps using no more than a dozen photos and the relevant movies, and the graph of the dialogue where Clytemnestra is begging Orestes for clemency in the *Choēphoroi*, both appear to be the same kind of animals. And that is a most important fact about outlines: that their photos, if we zoom in and examine their insides more microscopically, turn out to look again like graphs. So, story graphs are creatures that can be said to be *self-similar* in a general sort of way: the more detail we get, the more graph structure we get[35], a principle which we can express in its most basic form as:

The Self-similarity Principle.

An outline of a story is a story.

(The only level at which this breaks down, is the inside of an action sentence, where the constituting elements are not time dependent.)

Let me conclude this section by returning to our main task of creating a goal-friendly language for stories, and saying that the language of photos and

[35] I stress the fact that I do not mean that graphs at higher levels of analysis are necessarily *homeomorphic* to the rougher ones: just that they are the same types of beings.

movies is ideally suited to dealing with issues of intentions, goals and plans. In this, a goal is a far-away photo of a movie, which a character in a much earlier photo of the story graph can "see", in the sense in which we described in Section 1, a long time before he or she reaches it – or fails to do so. The different functions of *seeing* and *going-to* are ideally described in the photos-and-movies language. Of the four elements in a quest, the *searcher*, the *initial position* and *environment* can be talked about with reference to the story graph, with the *goal* or *destination*, usually set through some form of *seeing*, at some stage of the story's development, being the goal of a quest inside it.

Intentions and the resulting clear goals these create a story, seek to maximize causality in action. A leading character's intention pushes for strong causal movies or paths to the goals. For, obviously, the interesting and important goals require causal movies to be reached – if they could be reached by the mere passage of time, i.e. just the *t*-movies, they would not really be worth a story. ("A man was born, then grew up, went to school, married, had children, was infected by a virus and died" could hardly be considered a worthwhile story, in the usual sense.) In this context, the process of outlining is also extremely important. For it is this which allows us to have a perception of higher and lower, ultimate and intermediate goals. And it is also this, especially the principle of the Independence of Levels, which allows us to speak of obligatory and optional intermediate goals, of a higher or lower degree of determination in a process leading to a goal.

Winning a war as a result of winning and/or (on occasion) losing battles, is an apt metaphor for any kind of long-term struggle, and is naturally translatable into the language of graphs. But can a general plan all the battles in a campaign in advance?

And with this question, we move on to proofs.

4. Proofs

In Section 3 we described an application of a graph language to stories. We started from the cognitive primacy of the action sentence, or its more modern analogue of the comics panel; this gave us the dots of our graphs, which we called *photos*. Their arrows we named *movies*, and in these we saw the analogue of the vehicles of linear progress in the surface form of a story, these being combinations of the functions of temporal sequentiality ("and then") and causality ("and so"). One of the basic advantages of a graph language for stories is that it helps us define and examine *goals* and related concepts. Once we've define an *initial position* (a photo) and a *goal* to be reached (another photo, representing an action sentence in the future tense, like "I shall be king" or "I will marry Daisy", etc.), the process of going to it becomes a *quest*, in which the existence of sub-goals, partial goals, intermediate goals, etc., which can change dynamically during the process of the quest, is the infrastructure for a story.

Now, extending the photos-and-movies language to proofs comes quite naturally, for the staple of mathematics, already from the time of Euclid, has been propositions and deductions and so, in a sense, the present section could end on its very first page, with this statement:

"For *photos*, take mathematical *propositions*; for *movies*, *deductions*.

But by applying the photos-and-movies language to stories we've identified a number of basic concepts with cognitive analogues, and unless a graph language is also shown to have cognitive analogues for those in the new context of proof, the similarity will stay at the surface. Thus, to go deeper we must demonstrate how certain structural particularities of the photos-and-movies language as applied to stories translate to equivalent notions applying to proofs, and find convincing equivalences for:

- The space propositions and deductions occur in (the analogue of E-space).
- The relationship of non-linear to linear (the equivalent of going to/and from from "story graph" to "story")
- The nature of movies (the analogue for proofs of the two different functions, *t*- and *c*-, of story movies).
- The weighted movies, and how this applies to the two types in the new context.
- The concept of the linear and V- or Y-type joints.
- The concept of linearization.
- Outlines.
- The principles of Distance, Independence of Levels and Self-similarity.

Only when this is done, can we speak of the photos-and-movies language for proofs leading naturally to a structural equivalence in its application to stories, and thus showing forth deeper similarities. The methodology is similar to demonstrating that two graphs are homomorphic in mathematics. It is not sufficient for us to show how photos go to photos and movies to movies; we must also show that, via this translation, structures are mapped on similar structures.

4.1 The curious incident of the proof in the nighttime

Let us begin our discussion of logical proof with an astute insight by one of its most renowned advocates:

> "You see, Watson, it is not difficult to construct a series of inferences, each dependent upon its predecessor and each simple in itself. If after doing so, one simply knocks out all the central inferences and presents one's audience with the starting-point and the conclusion, one may produce a startling effect."[36]

Holmes's words bring to mind Alan Turing's definition of proof given in Section 2, as a "long march each step of which is trivial" (here: "trivial" is represented by "simple in itself"). However, the "startling effect" that the great Victorian sleuth refers to does not belong to the realm of *formal*, as to *folk logic*, i.e. the way in which the more rules of reasoning are applied in the context of everyday life. In fact, if we think of logic in these – as opposed to the *formal*, or *rigorous* – terms, the relationship of logical inference is more often than not qualified, or *weighted*. Unlike classical formal logic, which is the yes-or-no process par excellence, in its quotidian application the statement "A implies B" is very seldom taken as an absolute truth. Like causality in stories, it is often the additive effect of *more than one implications*, i.e. of various As being true, that can make B become true. Also – or is this the same thing? – implication in everyday life, as in the natural sciences, is more often than not an inductive, rather than a deductive, relationship.

Of course, in mathematics the concept of implication is much stronger. In fact, we are accustomed to think that in a mathematical context the statement "A implies B" has absolute value – and it certainly is much stronger than in any of the other sciences, coming as close as anything can to an absolute truth (always, of course, within the restrictions of the particular axiomatic system to which it belongs). And in a certain sense it *is* absolute: the word "implies" in the statements "that k is a non-prime integer implies that

---

[36] From the "Adventure of the Dancing Men", in Arthur Conan Doyle's *The Return of Sherlock Holmes*.

it can be decomposed into a product of primes in a unique way" or "that K is a subgroup of the finite group M implies that the number of elements of K divides the number of elements of M" obviously has a much stronger sense than the same word in the statements "that there are heavy clouds in the sky implies that it will rain today", or in "that A has lower cholesterol than B implies that A will live longer than B".

But even in mathematics, there are cases in which the word "implies" is not as strong as in these two examples. In fact, I will begin the discussion of the photos-and-movies language as applied to proofs with a question that many mathematicians might be tempted to immediately reply "yes" to: "Does the Taniyama-Shimura Conjecture imply Fermat's Last Theorem?"[37]

Now, in mathematics, the statement "*x implies y*" sometimes means "*if y is not true, then x is also not true*", and in this case, the verb "implies" in the statement "x implies y" is absolutely true, for otherwise it would lead to a case of *reductio ad absurdum*. So, obviously, when meant in this way, to say that "the Taniyama-Shimura Conjecture implies Fermat's Last Theorem" is absolutely true – in fact, it was the proof of this connection that Gerhardt Frey conjectured, and Ken Ribet proved, which opened the way to the proof[38]. But if "x implies y" is taken in another sense that is also frequent in mathematics, i.e. as synonymous to "*it is enough for x to be true for y also to be true*" – and the two senses are often mixed-up, even in the best circles, when we talk mathematics –, then the statement "the Taniyama-Shimura Conjecture implies Fermat's Last Theorem" is no more true than the statement "Gavrilo Princip's tuberculosis caused the burning of Dresden".

This may seem a shocking – even possibly a ludicrous – statement to some. But I am prepared to defend it. In fact, the first part of Section 4, which is about the application of the photos-and-movies language to proofs, can also be seen as an extended discussion of this position: that the word "implies" has two very different senses in mathematics, and that "Taniyama-Shimura implies Fermat" can be seen to be absolutely true in the first sense, but not so in the second.

4.2 Photos and movies in proofs

If sufficiently broken down into its elementary constituent parts, a proof should indeed consist of what Alan Turing calls "trivial steps". Yet trivial step by trivial step, mathematicians have proven the most incredibly deep and complex truths. How can this be?

---

[37] The question is meant to be asked to someone who knows of Andrew Wiles' famous proof.

[38] Frey conjectured that if Fermat's Last Theorem does not hold for a particular x, y and z, then it can be shown that the Taniyama-Shimura conjecture does not apply to them. Thus, exactly, "Taniyama-Shimura implies Fermat's Last Theorem in the sense that if Fermat's Last Theorem is *not* true, then neither is Taniyama-Shimura."

Let's call the places in a proof which these steps depart from and/or arrive at, its *photos*. Starting with the steps, as Turing does, you can also think of the photos as stones protruding from the surface of a deep, perilous lake, stones on which we can step in order to advance from one side of the pond to another, without drowning. These stones, or photos, are mathematical statements of fact, or *propositions*. Here is a sequence of some phrases from a proof out of Euclid, of theorem II.5 of the *Elements*. The string of non-underlined words in each phrase is a proposition, a *photo* in our language:

(8) …<u>therefore</u> the whole ΑΘ is equal to the gnomon ΜΝΞ.

(9) <u>But</u> ΑΘ is contained by ΑΔ, ΔΒ.

(10) <u>for</u> ΔΘ is equal to ΔΒ.

(11) <u>therefore</u> the gnomon ΜΝΞ, too, is equal to the (rectangle contained) by ΑΔ, ΔΒ.

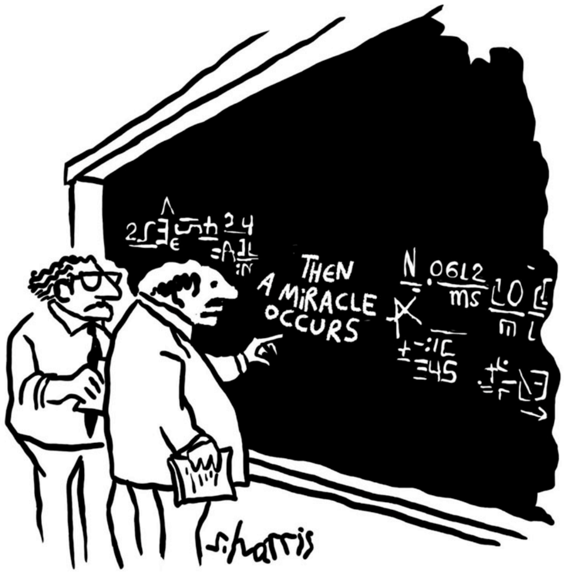
Photos can be trivial or profound, the simplest of statements like "1+1=2", "ABC is a right triangle", "17 is a prime number", "ΔΘ is equal to ΔΒ", to much more complex and deep observations, like "the primes are infinite" or "every finite group of odd order is solvable". (The last two are famous theorems, the first proved by Euclid and the second, much later, by Walter Feit and John Thompson.) Not every photo in mathematical discourse is *true* of course. But that is alright: to prove that the primes are infinite, Euclid also used a photo saying "the primes are finite" – which he later contradicted.

All the true photos taken together are what an extreme mathematical Platonist would call "the truths of mathematics". In the ethereal universe these entities share with the non-true ones – let's call it P-space, the "P" standing for "proposition" –, we may be excused to think that there isn't such a thing as time: Platonic truths are just *there*, and by their side also their negations. But if we want to move around in P-space, this will occur, like all movement, in *time* – or, depending on how you look at it, you can call this *space*, but space needs time to be traversed[39]. The total stasis of the set of all truth is upset by the movement indicated in phrases like "*x* implies *y*" or "let's write *a* as *b*", etc. These are the steps taking us from photo to photo, the *movies* of P-space, which of course is the analogue of E-space, when we are talking of proofs, instead of stories.

Now, a mathematician writing a proof will use both photos and movies. S/he will state facts (photos) but also connect them by various processes (movies), sometimes by simply ordering them in a particular way, at others

[39] For two discussions of the question of time in mathematics, in this sense, see G.E.R. Lloyd's "Mathematics and Narrative: an Aristotelian Perspective" (in the forthcoming *Mathematics and Narrative* book), and Barry Mazur, "On the Absence of Time in Mathematics (on his website).

deriving one from the other. (In fact, the underlined words in the lines from Euclid's proof, above, are the movies.) Of course, when they are presenting their proofs to their peers, mathematicians can skip many movies, for reasons of clarity and/or brevity. But when, in the process of doing this, they hear a sentence like the one told by the gentleman on the right in Sydney Harris's wonderful cartoon, they must be able to provide, on demand, the omitted parts:



"I think you should be more expicit here in step two."

In the process of trying to create a proof, mathematicians move inside a space which has a strong cognitive analogue, in its small neighborhoods. When they work on the small scale of a proof, they'll have a very clear sense of exploring by moving on, starting from one proposition and trying to go to another one, taking nearby roads (movies or paths thereof), rewriting something as something else, perhaps a decomposition or factorization, or maybe an expansion, substituting, playing combinatorial tricks with statements, connecting similar to similar. Of the huge, immensely complex being in which these are lodged, they need not think at that time. In this sense, though a mathematician will always know that his thoughts live inside the immensity of P-space (though he or she won't use that name for it), they only experience it, in the deductive process, bit by bit. In fact, the metaphorical magnifying glass that they are using for their high-precision work can only see a part of it at any one time, experiencing the riches of P-space photo by photo, as illuminated by proof. If a mathematician has a good memory and/or a good library, he or she can have access to very many proofs. But as we are time-breathing beings, we can only discover, read, teach, tell, and remember proofs *in time*. (The issues of imagination, intuition, non-rigorous thinking, and so on in mathematics, do not refer to this part of the process of proof, the time – so to speak – when the mathematician, chalk or pencil in hand, is working on the microstructure of a proof.)

Though T.S. Eliot is not speaking of proofs in “Burnt Norton”, I cannot say it better than he does:

> …Only in time can the moment in the rose-garden,
> The moment in the arbour where the rain beat,
> The moment in the draughty church at smokefall
> Be remembered; involved with past and future
> Only through time time is conquered.

The movies in a proof require time to come to life, to *move*, and a time-fuelled machine, such as a mathematician, to bring them to being. Only through time can proofs become real.

4.3 On links in P-space and the nature of deductive movies

Proofs are recorded and are read linearly, word after word after word, in time. But are they really linear?

Imagine you are a minute rational being living in the (locally) one-dimensional world of movies and photos. Say, for example, you are some sort of mathematical ant in P-space, able to do two things: to sit on photos, and to walk on movies in the direction in which their arrow is pointing. You also possess some amount of memory, and vision, enabling you to see, say, one photo ahead every time, i.e. to allow you to see, when you are standing on a photo, the photos that are connected to the one you are standing on by exactly one movie; and you also have a lexicon, which can supply you, on demand, with the meaning of terms you haven’t seen before. The basic work mathematical ants have to do is to move around and “discover” proofs. The way to do this, in ant-world, is to start from photos which some higher deity in the mathematical universe tells them are called “true”, and then walk from there down any path they want, as long as every movie in it has the number ☺ written on it, a number which our ants can recognize. (Let us call, in verbal language, each movie marked with a ☺, a “*fully legitimate deductive step*”.)

Now, let us assume that I, who happen to be one of the higher deities, place you, the ant, on photo (0), and give you the information – I am a totally dependable source, by the way – that (0) represents a true statement. This is it:

(0) ABC is a right triangle.

From there, you look around and choose to walk down a movie (marked ☺) which takes you to another photo you recognize:

(1) $a^2 + b^2 = c^2$

Being a mathematical ant, you know – or, if it doesn't ring a bell you can always look it up in your lexicon – that the letters *a*, *b* and *c* stand for the sides of triangle ABC referred-to in (0) and that (1) is a famous statement, which has the name "Pythagorean Theorem". (Incidentally, you can also recognize a famous theorem from the fact that it has an unusually large amount of traffic going through it.) From (1) you embark on a movie, also marked ☺ – you may happen to remember that this particular movie is a type called a "rewrite rule", but of course as long as traversing it is concerned it's just another movie – which takes you to:

(2) $a^2 + b^2/c^2 = 1$

From there, to another:

(3) $a^2/c^2 + b^2/c^2 = 1$

Then to another:

(4) $(a/c)^2 + (b/c)^2 = 1$

All these marked with ☺. Now, you look up the last photo in your dictionary and see that, if *a*, *b*, and *c* refer to the sides of a right triangle, you can rewrite *a*/*c* and *b*/*c* with other symbols, so you do, which process is in fact also a ☺-marked movie, and you end your stroll at:

(5) $sin^2 \mathrm{A} + cos^2 \mathrm{B} = 1$

Now, one of your duties as a mathematical ant is to record all your travels. So you will write down this little itinerary, happy again double-checked that all the movies have ☺ on them (you are a pretty obsessive little ant). You can write down your journy as:

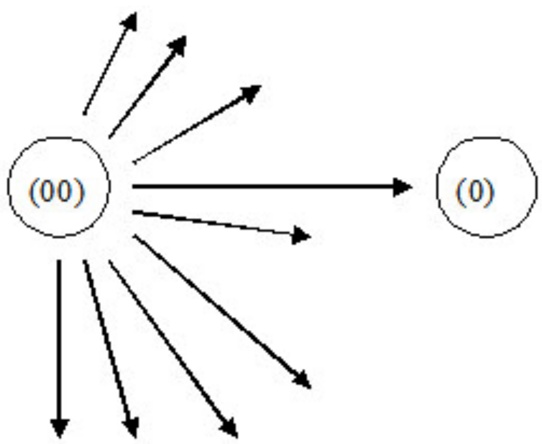
(0) → (1) → (2) → (3) → (4) → (5)

In the evening, when you sit and have a barley drink with your fellow mathematical ants, you will recount to them your little journey. They will probably be quite impressed – especially if they have not gone that particular way themselves, and even more especially if you do the Holmes trick, and say, in a rather nonchalant way: “Hey, did you know that in right triangles the square of the sine plus the square of the cosine of an angle is 1?” If, after being startled, they ask you to *prove* this, you will just recount your journey, maybe bore them a bit, but also convince them. (The reason why they may be bored is that (2) and (3) and (4) are not very fun locations.) But (5) they will like a lot, they may call it a “cute little theorem” and congratulate you on the fact that you have produced a proof, a path from (1), i.e. something known, to something *both* previously-unknown and very nice.

But now assume another mathematical ant, who has heard your story, wants to retrace your nice little path. Because I am a slightly mischievous ant god, a trickster of sorts, I will place the second ant on photo (00) which is just one movie away from (0) and watch to see what happens. This is it:

(00) ABC is a triangle.

Now, when the ant looks out from (00), it can see that there are a number of movies starting from (00), as in the diagram below[40]:



[40] Actually, in the reality of P-space it’s much much more complicated than this – with many more incoming and outgoing movies from- and to- both photos, but let’s just leave it at that, for now.

But that will be alright as, being able to see one photo ahead, the new ant will be able to see (0) from (00) and so go there, to start on his way to the new theorem. And once he gets there, assuming he has a pretty good memory, he will be able to repeat his friend's journey, and give them his proof as:

$$(00) \rightarrow (0) \rightarrow (1) \rightarrow (2) \rightarrow (3) \rightarrow (4) \rightarrow (5)$$

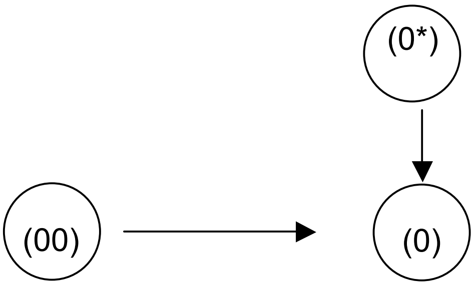
But if that evening, at the barley-drink bar, he starts the story to his friends from (00), the fact that he walked down the movie (00) → (0) will not be enough to convince them that (0) is a right triangle. And when they ask him, "Hey, how do you know that the triangle you are speaking about is right?", he won't be able to give them a convincing answer. The reason for this, you see is that this ant has not been careful, and did not notice that (00) → (0) was not marked with a ☺ but with a 😐 and every mathematical ant knows that when you walk down a 😐 movie from X to Y, you are not allowed to describe this to your friends as "X *implies* Y". We can rephrase this (00) → (0) being a 😐-type movie in people language, as: "That a shape is a *triangle* (00) does *not* imply that it is a *right* triangle (0)." And, indeed, it does not.

But now assume that at (0) our new ant meets a friend who has just arrived there via another incoming movie from photo (0*), which is this:

(0*) C is a right angle.

Now, though (0*) → (0) is also marked with a 😐, if the two ants combine their accounts and *tell both* to their friends, this will in fact convince the myrmecian powers that be that (0) is, in fact, *proven* to be true. But we must stress here that *both* (00) and (0*) are necessary for the proof. You see, (00) of itself just tells us ABC is a triangle; and (0*) of itself just tells us that an angle C is right – but it could be an angle of a rectangle and not a triangle.

But the two together, will be accepted as proof, not just to the ants, but also to me and my fellow deities: for you see, as we get the input from our faithful (the ants), we combine it and draw it on pieces of paper: we deities, being superior beings, live in a two-dimensional space, and converse with other deities who also possess two-dimensional knowledge and can also perceive diagrams. The following diagram, which I can construct from the reports of my two ant scouts, is enough to convince my peers – once they accept (00) and (0*) as true, that the ABC is indeed a right triangle:



But it’s interesting to note, that though the above diagram will be appealing to the other deities’ visual sense, as it is to mine, if they want to check that it is in fact a good proof, they will have to know the nature of the two movies (i.e. whether they are either ☺-type or 😐-type and not the bad non-deductive movies of ☹-type) and go over them, *first* (0*) → (0) and *then* (00) → (0) -- or the other way round, of course, but not simultaneously: *only through time*. This process I’ll call linearization (of a non-linear fact).

* * * * *

Now, let’s forget ants, and try to see how what we saw in their case about the two kinds of good movies in P-space, i.e. those marked with a ☺ and those marked with a 😐, can apply to the world of human mathematicians. In fact, let’s look at how V-shaped junctions such as the above enter a proof. We’ll look at the structure of a rather simple – but greatly important – very old proof, and at a small bit of a very famous and very complex one.

First, the very old one: this is a variation on Euclid’s glorious proof of the infinity of prime numbers, expressed in words, i.e. *linearly*, in a way which is rigorous enough to convince a mathematician but not too formal – it’s clear that some of the photos in it can be broken down to even more elementary ones, or even simple subgraphs of photos and simple deductions, but let’s not be total pedants and make this as user-friendly as possible.

The underlined words indicate the traces of movies:

(1) Assume the primes are finite, written as the ordered set P = (p_1, p_2, p_3… p_n) where p_n is the largest prime.

(2) Construct the number p_1 x p_2 x p_3 x… p_n + 1 and call this *k*.

(3) *k* is either prime or non-prime.

(4) <u>If</u> *k* is prime, <u>this leads to</u> a contradiction with (1), <u>as</u> *k* is obviously larger than p_n, the largest number in the set P of all primes, <u>and thus</u> cannot be a member of P, <u>which, however,</u> we defined as the (finite) set of all primes.

(5) If *k* is not prime, then there is a prime which divides it – call this *p**.

(6) *p** is not be a member of P, <u>for</u> k divided by *p** leaves a remainder of 1.

(7) (6) contradicts (1) – which says that P contains all primes.

(8) The primes are infinite.

Now, though the sequence (1) to (8) represents the proof linearly, word after word after word, its form in P-space is not a path nor a *tree* but, as a photo may connect to more than one subsequent photo, a *dag* (see definition, bottom of page 40) which is *at least as big* as the following diagram, where (9) is the Rule of the Excluded Middle, (10) is a corollary of the Fundamental Theorem of Arithmetic[41], and (11) the Distributive Law for Division[42]:

$$
\begin{tikzcd}
& (9) \arrow[dr] & & & & & & \\
(1) \arrow[r] & (2) \arrow[r] & (3) \arrow[r] & (4) \arrow[r] & (5) \arrow[r] & (6) \arrow[r] & (7) \arrow[r] & (8) \\
& & & (10) \arrow[ur] & (11) \arrow[ur] & & &
\end{tikzcd}
$$

Now moving to the top end of modern mathematical sophistication, let's look at a few lines from Andrew Wiles' extremely famous proof of the Taniyama-Shimura Conjecture, which, along with Ken Ribet's proof already mentioned, that Taniyama-Shimura implies Fermat, clinched the matter of the Last Theorem. It is, I think, extremely interesting that to understand that the proof of this superstar theorem, like the proof of the humblest little proposition, also consists of photos and movies, you do not need to know the first thing about algebraic number theory, just the common English words Wiles uses, connecting his – to a layperson – totally abstruse language.

I have highlighted the movies in green, and the photos in pink, and I'm sure even the most non-mathematical reader can see why each is which quite easily[43]:

---

[41] That every non-prime number can be decomposed into a product of primes in a unique way (but for the order of the factors).

[42] That $a+b/c = a/c + b/c$

[43] The black part is from page 470 of Andrew Wiles's "Modular forms and Fermat's Last Theorem", *Annals of Mathematics,* **141** (1995), 443-551.

Proof. We first observe that $\mathrm{pr}_n(H^1_F(\mathbf{Q}_p, T)) \subset H^1_F(\mathbf{Q}_p, T/\lambda^n)$. Now from the construction we may identify T/λ^n with V_{λ^n}. A result of Bloch-Kato ([BK, Prop. 3.8]) says that $H^1_F(\mathbf{Q}_p, \mathcal{V})$ and $H^1_F(\mathbf{Q}_p, \mathcal{V}^*)$ are orthogonal complements under Tate local duality. It follows formally that $H^1_F(\mathbf{Q}_p, V^*_{\lambda^n})$ and $\mathrm{pr}_n(H^1_F(\mathbf{Q}_p, T))$ are orthogonal complements, so to prove the proposition it is enough to show that

$$\#H^1_F(\mathbf{Q}_p, V^*_{\lambda^n})\#H^1_F(\mathbf{Q}_p, V_{\lambda^n}) = \#H^1(\mathbf{Q}_p, V_{\lambda^n}). \tag{1.12}$$

Now if $r = \dim_K H^1_F(\mathbf{Q}_p, \mathcal{V})$ and $s = \dim_K H^1_F(\mathbf{Q}_p, \mathcal{V}^*)$ then

$$r + s = \dim_K H^0(\mathbf{Q}_p, \mathcal{V}) + \dim_K H^0(\mathbf{Q}_p, \mathcal{V}^*) + \dim_K \mathcal{V}. \tag{1.13}$$

From the definition,

$$\#H^1_F(\mathbf{Q}_p, V_{\lambda^n}) = \#(\mathcal{O}/\lambda^n)^r \cdot \#\ker\{H^1(\mathbf{Q}_p, V_{\lambda^n}) \to H^1(\mathbf{Q}_p, V)\}. \tag{1.14}$$

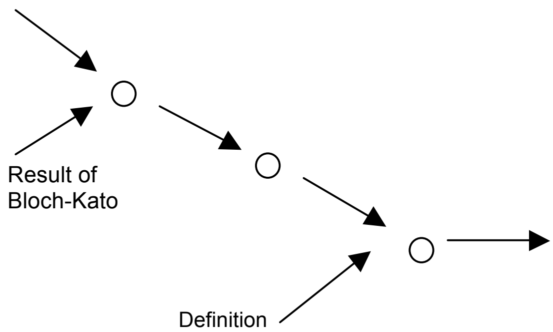
The second factor is equal to $\#\{V(\mathbf{Q}_p)/\lambda^n V(\mathbf{Q}_p)\}$. When we write $V(\mathbf{Q}_p)^{\mathrm{div}}$ for the maximal divisible subgroup of $V(\mathbf{Q}_p)$ this is the same as

$$\begin{aligned}\#(V(\mathbf{Q}_p)/V(\mathbf{Q}_p)^{\mathrm{div}})/\lambda^n &= \#(V(\mathbf{Q}_p)/V(\mathbf{Q}_p)^{\mathrm{div}})_{\lambda^n} \\ &= \#V(\mathbf{Q}_p)_{\lambda^n}/\#(V(\mathbf{Q}_p)^{\mathrm{div}})_{\lambda^n}.\end{aligned}$$

Combining this with (1.14) gives

$$\begin{aligned}\#H^1_F(\mathbf{Q}_p, V_{\lambda^n}) = \#(\mathcal{O}/\lambda^n)^r \\ \cdot \#H^0(\mathbf{Q}_p, V_{\lambda^n})/\#(\mathcal{O}/\lambda^n)^{\dim_K H^0(\mathbf{Q}_p, \mathcal{V})}.\end{aligned} \tag{1.15}$$

As you see, there are at least two points in this short snippet of the proof (which represents something like a one-two hundredth of its total length) where Wiles appeals to more than one photo, i.e. uses more than one movie, as he continues to the next in the linear sequence. One of these is indicated by the phrase “a result of Bloch-Kato” and the other by “from the definition”. Thus, already in this tiny fragment of the proof we have something like this:



Some of the non-linear (2 or more-to-1) V-junctions in the graph of the full proof, like the two upwardly directed ones in this diagram – there are obviously many more throughout the paper –, are indicated in the proof-as-published (linear form) through verbal markers, verbs like “observe” or "follows", and/or prepositions and conjunctions. But even for the case of those junctions which do not leave a visible verbal trace of their non-linearity in the proof-as-published, there is no doubt than a strong expert in the field, if given sufficient time, can detect their invisible operation.

Given the ubiquitous non-linearity of proofs, from the microstructure of junctions (the 'molecular level') upwards, we see that the innocuous-looking statement "A implies B", can have two meanings, as they are given to us, in the simplest case, by the two following diagrams:

(10) Type-1

$$A \longrightarrow B$$

(11) Type-2

$$\begin{tikzcd} A \arrow[dr] & \\ & B \\ A^* \arrow[ur] & \end{tikzcd}$$

The movie in type-1 diagrams is always ☺. The movies in type-2, are always 😐 -- for if there is even one ☺ movie going from a true A to a B, this is enough to consider B proven. Thus, though situations as in diagram 12 do exist, in the context of a proof the lower movie is superfluous:

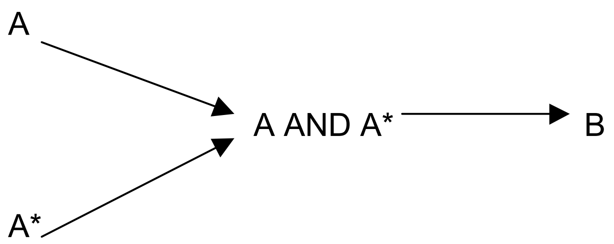
(12)

$$\begin{tikzcd} A \arrow[dr, "☺"] & \\ & B \\ A^* \arrow[ur, "😐"'] & \end{tikzcd}$$

There is a very powerful corollary to the type-1 and type-2 diagrams, which gives us one of the most potentially profound insights into the relationship of stories and proofs – I hope its validity will become more apparent as we progress, especially in the historical Section 5. This is it: using the photos-and-movies language, and given the two dimensional sight of the myrmecian deities, and not their humble believers, to understand whether a movie is type-1 or type-2 *we do not really need to verbally brand a movie as anything*, no, nor even by our cute little symbols ☺ and 😐, and especially not the ambiguous word "implies" or any variations of things like "substitute this for that", "rewrite as", etc. All we need is to look at the arrows. All the important information is there, in the graph structure of a minimal (and sufficient, of course) statement of the proof, with the most important distinction in the graph of a proof being that between movies which can act alone inside the graph, i.e. the type-1 movies – such is, for example, the passage from (0) to (1) in the proof of the simple trigonometric theorem we referred to earlier, on pages 55 and 56 – or the movie from (1) to (2) in the case of the proof of the infinity of primes, on page 59.

But type-2 movies can never work alone. In these, the implication is weaker, as it were – a fact which we can also describe with the fact that movie is weighted, with a number strictly between 0 and 1. An example of a type-2 movie is thus (00) → (0) (page 57) – incidentally, it’s interesting to note here that (0) → (00) is type-1. Though in type-2 movies, B can be said to “logically derivable” from A in some way or other – perhaps a way only understandable in non-formal terms –, and thus in a certain sense also “A implies B” it, if we want “implies” to acquire the strong meaning of “is a necessary consequence of”, we *cannot* reduce diagram 11 on the previous page to “A implies B”. For a rigorous transcription in linear/verbal form we have to say “A AND A* imply B”, where AND is the usual Boolean connective known to every internet user.

One could attempt to translate a type-2 situation into type-1, as in the diagram below.



But, though it succeeds in replacing the incoming movie on B, from type-2 to type-1, the problem is pushed further back, and thus the two movies going from A and A* to A AND A* are necessarily type-2.

After having gone through this, we can now qualify our shocking statement in 4.1, i.e. that “saying that ‘the Taniyama-Shimura Conjecture implies Fermat’s Last Theorem’ is a true as saying that ‘Gavrilo Princip’s tuberculosis caused the bombing of Dresden’”. Now we can state that, obviously, the statement that the “Taniyama-Shimura Conjecture implies Fermat’s Last Theorem” can only be said to be true if the word “implies” in it is meant in the weak sense, i.e. as a type-2 movie. To make it type-1, we would have to construct a previous photo to Fermat’s Last Theorem, as in diagram (12), above, saying : “The Taniyama-Shimura Conjecture AND P_1 AND P_2 AND P_3 … AND P_n imply Fermat’s Last Theorem”, where P_1, P_2, P_3…, P_n

would be all the results deemed necessary, first by Ribet, and then by Wiles, in the arguments presented in their papers. And on this photo, of course, all these P_1, P_2, P_3…, P_n would have to converge, to make it valid. (In fact, we know at least one of the P_i's: it is the "result of Bloch-Kato" referred to by Wiles on page 470 of his paper, and here quoted on our page 61.) And, of course, all of these incoming movies are type-2[44]. It is only this huge (the *n* of the highest P_n is very high) amount of convergence on it, that actually proves Fermat's Last Theorem. But that is natural: if it was less populated, it wouldn't take a man of Wiles's huge talent and iron determination and will-power seven long years and 108 pages of extremely difficult mathematics to prove it!

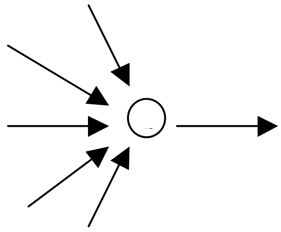
Yet, despite the strong difference between type-1 and type-2 movies in proof, mathematicians in their discourse often use the word "implies", or "therefore", or "thus", or "so", and some related terms, without qualifying which of the two senses they are referring to. Though it may be argued that this is done so more in informal than in formal discourse, it is the informal ways which are much better indicators to the underlying, cognitive infrastructure of proof.[45]

* * * * *

So, at least half of the non-linearity of proofs, giving us their dag structure like graph structure (on occasion we can depart from the dag structure via loops, when some process must be repeated), with the theorem proven as the *root*, and the various necessary photos used in the proof (definitions, axioms, other people's theorems and lemmas proven on the way to the proof) gradually converging on it, is a result of type-2 causality, the fact that the incoming movies on photos in P-space as a rule look like this:

[44] To be true, a statement containing the Boolean AND requires, by definition, all the elements connected with AND to be true.

[45] But here is an example of it being stated seriously in the most respectable of mathematical circles. The exact phrase "the Taniyama-Shimura Conjecture implies Fermat's Last Theorem", without caveats or qualifications, appears in the piece describing Wiles's seminal article in the *Mathematical Reviews* published by the American Mathematical Society (see it reprinted in its 'Selected Reviews' in the New Series of the *Bulletin of the American Mathematical Society*, Volume 43, Number 3, July 2006, p. 409).



But P-space is not really all trees. We don't just have the many-to-one but the one-to-many joints, making for the dag structure – including possibly the allowing for the odd loop? Look at our little ant in the middle of page 57, looking out of (00), seeing (0) *and many other photos?* Or, what about if we stand on photo (0), that "ABC is a right triangle" and look the other way from (00)[46].

Standing at (0), an ant would see another huge bunch of movies opening up, and leading to all sorts of directions. Here are some of the photos these would lead to:

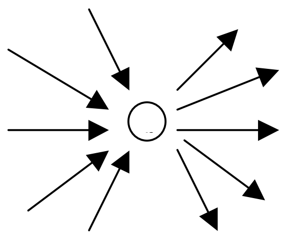
(1_1) A+B = C

(1_2) A < C and B < C

(1_3) The construction resulting from gluing ABC and an exact copy of it flipped in a certain way across the side AB, will be a rectangle.

And of course it could also see photo (1) in our trigonometric proof (page 56), i.e. the Pythagorean Theorem. And this is pretty much the situation of every photo of P-space: *not only does it have many movies coming in to it, but also many movies going out from it.*

And this means that photos in P-space don't look like trees but veritable sea-urchins:



[46] In a fuller version of the subgraph, there must appear also the movie *returning* from (0) to (00), which as we mentioned is type-1 and not type-2. This of course means that "ABC is a right triangle" absolutely implies that "ABC is a triangle", a Kantian analytic statement if ever there was one.

And this is the reason why mathematical ants must have good memories, with which to reconstruct proofs – once discovered by one of their peers and recounted to them –, or else they would be totally hopeless in finding them once again. Because if the ants had to advance in a random way, even in the simple situation of our trigonometric proof, the linear recording of which at the bottom of page 56 has only four movies, from the Pythagorean Theorem onwards, assuming that each photo has only 5 outgoing movies in P-space – and this is a huge underestimation –, the probability of reaching (5) from (1) in just five steps, would be on the order of 0.0003.

So, though a final proof-as-published may well look like a dag (maybe with loops), the process of discovering of a new one – unless the mathematician is uncannily lucky – is bound to be a much more complex subgraph of P-space.

4.4 Proof-as-discovered, proof-as-published

A published proof represents the epitome of mathematical *savoir faire*. Rigor, clarity and completeness are absolutely required of it, and economy is also appreciated, though never at the expense of the first three qualities. At the heart of a paper published in a serious mathematical journal is one or more theorems. The paper's purpose is essentially rhetorical, aimed as it is at *convincing* its readers of the truth of its assertion, the theorem. But as the readers are the mathematicians who are experts in the field to which the theorem belongs, the writer can only use the types of arguments that these folk find convincing, as determined by peer review. The *lege artis* guiding the practice of the mathematical clan is here represented in its purest form.

Now, according to ancient theory, the parts of a rhetorical speech are three: *ethos* (i.e. an argument on why the particular speaker is ideally suited to expound on this matter), *pathos* (i.e. an appeal to the audience's feelings), and *logos* (i.e. reason). In a proof, the first is actually quite strong, though never expressly stated, for this would be against the *savoir faire* – for example, in the extract from Wiles's proof that we presented, the fact that Wiles says something is a result of Bloch-Kato guarantees to the average mathematical (expert) reader that he can take it for granted, for all three, Bloch, Kato and Wiles are honorable men – and the veracity of the result is further guaranteed by the *ethos* of its journal of publication. (Of course, if a mathematician is caught once too often making public statements which have not passed the crucible of peer review, s/he may lose a lot of that credibility.) The second element of rhetoric, *pathos* is absent from mathematical arguments: sentences like "you must believe that manifolds A and B are isomorphic because it is honorable", or "because otherwise I will not get my tenure and my grandfather will die of shame" are strictly taboo.) But *logos*, the third element, and this in its most powerful, and distilled form, predominates in proof-as-published.

Now, proof-as-published is a goal-directed activity *par excellence*. The intentions of a mathematician writing a proof for publication will be clear from a paper's beginning to the reader, either in a statement very near the top of the paper, or even in the title (or else the abstract). The mathematician will indicate the result(s) s/he wants to prove, s/he may briefly describe the rationale of the strategy to be adopted, or move directly into its presentation. In any case, a proof-as-published is a perfect example of a completely goal-dominated and goal-determined process. If we make a diagram of the photos and movies of even the tiniest part of it, we will see even in its geometry that everything is dominated by intention, the goal, the thing-to-be-proven even in this smallest part of it almost pulling everything towards it, all the previous photos shooting their arrows in its direction. Everything will be focused on achieving the intended result, the QED. All the movies will move towards the final photo, all the little photos on the way will be optimally necessary for the task – the ultimate goal will be *the root of the dag*, i.e. the final photo from which no other departs: the theorem.

But we must remember here that a proof-as-published is only a highly distilled form of the proof-as-discovered. And if the proof-as-published is clean, and as concise as possible, and non-ambiguous, and all sorts of good things, the proof-as-discovered can be messy and repetitive, and vague, and full of gaps (that are eventually filled) and often boring and frustrating and many other bad things…

Now, though these "bad" elements of proofs-as-discovered are ubiquitous in the oral environment of mathematics, both in the case of mathematicians actually conducting their research in restricted or more extended forms of collaboration, and in the case of their talking about their successes and the failures to their peers, the are practically absent from the written, published culture – in fact, as a rule you only meet them in some review articles also containing historical asides, or of course in the recent flux of books describing the stories of famous problems, like Fermat's Last Theorem, the Riemann Hypothesis and, very recently, the Poincaré Conjecture[47].

---

[47] A very rare example of a research paper announcing a proof which also contains a description of the history of its author's attempts at the proof, and thus elements of the subgraph of P-space that concerns the proof-as-discovered, is Wiles's article of his proof of Taniyama-Shimura, quoted above. In fact, it begins with an abstract which is a mini story of a type definitely *not* usually found in mathematical papers – it even includes a picture of the author as a child! --- and then goes on to give an 11-page account of the progress of the attempt of the proof, including references to some of the "bad" things, before it gets into the real mathematics of it, the *actual* proof. But this is a rare exception to mathematical *savoir faire*, obviously attributable to a combination of the problem's star-quality, and the immense

Clearly, the proof-as-discovered will not seem to be, in general, as clearly focused on the goal as the proof-as-published, though a lot of it may be. Yet, the proof-as-discovered can also be mapped in P-space – *in fact, the proof-as-published will be a subgraph of it.* The proof-as-discovered will contain, in addition to the good mathematics that will appear in the proof-as-published, also paths and dags which will be abandoned, subgraphs of P-space created not by iron logic and strict goal-orientation – as every part of the proof-as-published will appear to be – but caused also by chance, contingency, mistakes, misjudgment, trial and error, i.e. the "bad" stuff.

It's pertinent to note here that many of the elements in the proof-as-discovered that will be excised to create the proof-as-published, will be the elements that appear most story-like to an outsider. In fact, if one wants to write the story of a mathematical proof for a non-mathematical audience, as Simon Singh did in his bestseller *Fermat's Last Theorem*, all the chance, the contingency, the errors, the cul-de-sacs, plus a lot more will not just be included but given prominence, to give the story added human interest[48].

But even without those, for a mathematical reader the intellectual adventure itself, i.e. the graph tracing the attempts at proof, will be exciting in a way which can easily be called *narrative*: the narrative structure will be there – after all, if translated in photos-and-movies language, and we forget the content of the photos and the labels on the movies, the proof-as-discovered will look exactly like a story – and the emotions, the *pathos*, will be added by the reader. A mathematician writing the story of his/her discovery for an audience of some mathematical sophistication does not need to write "when I discovered that method X wasn't working I cried my head off, and seriously contemplated suicide", or "as I was watching a glorious sunset in Majorca, sipping a cold Martini, a fabulous new idea came to me on how Y's Lemma could be made to work", to convey a sense of the intellectual drama. Assuming the detail is pertinent to the quest, the author sincere, and the reader well-informed, the account of the journey towards discovery can be simply mathematical – and yet convey all the sense of the adventure. And it may also be particularly useful to mathematicians, not just for its inspirational (emotional and/or moral) qualities, but because it will allow cognitive insights into methodology, it will show them how the mathematician *thought* in order to get to his proof in a much clearer way than in the proof-as-published, where a lot of the quest procedure may have been excised, for reasons of clarity and conciseness. But this interest in the mathematical quest is interesting

publicity – at least by mathematical standards – generated by Wiles's early announcement of a proof, the subsequent discovery of a gap, and the final triumph.

[48] One of my favourite quotations is by the British masterspy Kim Philby, quoted in Genrikh Borovik's *The Philby Files*, explaining why he does not consider any spy thrillers, even those by John Le Carré, at all realistic: "…They do not include the human factor in their plots, that is *chance* and *error*." (*My italics.*)

precisely because the progress can be described in the language of photos-and-movies, *i.e. the exact same language with which stories are made.*

What's also pertinent to our argument is that, apart of the appreciation of its *logos*, the expert reader will also react to the proof-as-published in ways which the non-mathematician can feel only when exposed to the narrative mode: suspense, anticipation, surprise, disappointment and joy are very frequent reactions of a mathematician as s/he is reading the proof – and I mean here such feelings resulting from the proof itself, and not its relevance to the reader's own ambitions (i.e., I do not include that the reader may feel joy in the proof of a certain theorem because it will give him/her a tool for his/her own proof, or fear that the writer has discovered something that the reader is about to publish as new.)

4.5 Seeing, plans and goals: moving around in P-space

> "Would you tell me, please, which way I ought to go from here?"
> "That depends a good deal on where you want to get to," said the Cat.
> "I don't much care where," said Alice.
> "Then it doesn't matter which way you go," said the Cat.
>
> Lewis Carroll, *Alice's Adventures in Wonderland*

> "Hullo," said Holmes suddenly. "What's this?" It was a small wax vesta (*match*), half burned, which was so coated with mud that it looked at first like a little chip of wood.
>
> "I cannot think how I came to overlook it," said the Inspector, annoyed.
>
> "It was invisible, buried in the mud," said Holmes. "I only saw it because I was looking for it."
>
> "What," said the Inspector. "You were expecting to find it?"
>
> "I thought it not unlikely," said Holmes.
>
> Arthur Conan Doyle, "Silver Blaze"

Both of these quotations date from the second half of the nineteenth century. The author of the first was a specialist in logic, a disciple of George Boole, the man who gave mathematical form to what he called "the laws of thought" – in fact the very same language that later made it possible for mathematicians to enter P-space and create a science of metamathematics. The second quotation is from a doctor-turned-storyteller, who created the most prototypical literary character of the modern scientist as detached reasoner. In both, lurks the image of the world of knowledge as a dark labyrinth of possibility, which needs criteria (goals, desires, preferences, etc.) in order to be successfully traversed. Indeed, a mathematician thrown inside the jungle of complexity of P-space can only hope to achieve results by[49]:

A. Thinking of a proof not just in details but also in outlines.

B. Making guesses.

(Interestingly, both methods have been discussed in Section 3 in the context of stories.)

Now, regarding A, the methodology of outlining graphs of proofs can be taken to be the same as that of outlining stories, and the same restrictions and rules apply. In fact, this is a good time to underline what the reader has certainly already noticed: that there is a direct analogy, on the one hand between *t*-movies in story graphs and type-1 movies in proof graphs, and on the other *c*-movies in story graphs and type-2 movies in proof graphs. And through this analogy, the three principles, i.e. Distance, Independence of Levels and Self-similarity hold in a proof graph as well as a story graph. It is with the aid of outlines, and these principles, that a mathematician can understand and structure a search on the large scale, break it down, and approach problems piecemeal.

Let us now concentrate on case B, the second of the factors making a mathematician's orientation in P-space simpler: making guesses or, in another terminology, having *goals*. Goals are the main weapon mathematicians have in navigating the formless immensity of P-space. In fact, it is goals that make P-space something other than a meaningless inferno, a Borgesian "Library of Babel", where all exists and nothing matters. And the difference between *guessing* and *proving* in the context of proof is exactly analogous to the difference between *seeing* and *going to*, as we saw it operate in stories, in Section 3.

[49] Of course, another way used by both storytellers and mathematicians to traverse their infinite spaces, is relying on well-known bags of tricks, repertories and patterns. The reading, on the one hand, of books like the *Poetics*, the nineteenth-century Georges Polti and his taxonomies of the "thirty-six story situations", the work of Propp and other scholars who tried to understand the patterns in narratives, Northrop Frye's "archetypes of literature" or John Cawelti's studies of the typology of various genres, to the how-to books of Lajos Egri, Syd Field, Robert McGee, and similar ones, and on the other George Pólya's famous *How to solve it* – in which, incidentally, a dominant part is played by the notion of a *plan* –, and quite a few recent books in that vein, most notably among them the one by Fields Medallist Terence Tao, are testimony to the fact that practitioners of both narrative and mathematics rely on patterns to usefully reduce choice, and thus, also, complexity. And, of course, probably most practitioners of these fields learn these patterns and techniques to find and apply them *without* reading any of these books: in fact, studying and reading and learning and expertise-acquisition in both writing and proving is about learning to see patterns, and use them.

Keith Devlin has written of the famous coffee-fuelled theorem-proving machine Paul Erdös that his "greatest strength was as a problem solver who had a rare ability to ask questions that had just the right degree of difficulty". This astute observation makes Erdös a king, not just among problem solvers, but among goal setters. But not all mathematicians need be like him. Not every mathematical journey of adventure need have an ending similar to Homer's *Odyssey*. Ithaca – the goal originally planned-for – may indeed be reached but, upon arrival, found to be less interesting than had been imagined, whereupon the researcher, like Tennyson's or Kazantzakis's Odysseus, will depart again to strive and seek anew, maybe to find more interesting things. Or, as in Cavafy's poem of that name, Ithaca may be no more than an excuse to set out on the journey, the things discovered on the way turning out to be much more interesting than the actual destination[50].

The aim to which a mathematical investigation is oriented, as a rule first enters the potential prover's mind as a photo representing a conjecture, either his/hers or not (e.g. the Poincaré Conjecture); but it can also be a larger, structured set of conjectures (e.g. the so-callled Langlands Program) or, more generally a research program (e.g. Gorenstein's strategy for the classification of finite simple groups), consisting not just of photos but proof graphs, outlining a general plan of attack. Of course, it can also be – as a matter of fact, it usually is – something much smaller and/or modest.

The photos-and-movies language can obviously be used to depict the proof graph of a proof-already-discovered, both in its actual structure and/or its position in a wider, surrounding subgraph of P-space. This can be done at any level of detail, all the way down to the very minute, molecular deductive steps, such as "rewrite a^2 as a x a". A graph of this kind, almost always an inverted tree (in more complicated cases, a dag) with the theorem to be proven at its root, appears in the proof-as-published as a seamless miracle of goal-orientation: everything is ideally placed, and finely tuned to conclude at the desired goal (the root of the dag) as optimally as possible. But the sheer economy of this graph tends to hide the logical adventure behind it. The demands of rigorous exposition antagonize the realities of cognitive adventure[51].

---

[50] It is definitely in this spirit that David Hilbert's answered when he was asked why he had not attempted to prove Fermat's Last Theorem: "Why should I try to kill the goose that lays the golden eggs?" (After all, isn't the Taniyama-Shimura Conjecture much more interesting mathematically than the assertion in Fermat's Last Theorem?)

[51] It's very interesting, in this context, that the great mathematician Laurent Schwartz, one of the earlier members of the Nicholas Bourbaki group, speaks with disdain in his autobiography, *Un mathématicien aux prises avec le siècle* (1967), of the *novelistic* (his term) books with which his generation had to learn mathematics, and to which they reacted with the *Euclidean* (again, Scwhartz's choice of term) methodology of the books of the Bourbaki canon. The books of the *Collection Borel*, which he criticizes, where in fact much closer to the

The graph of the quest of a proof is much more interesting. In fact, the *developing* graph of the quest of a proof, i.e. the graph of the quest as it grows in time, is a graph that comes face-to-face with all the vast complexity of P-space. But what is particularly interesting is that the general goal-orientation in the published proof will appear much more clearly if we look at a diagram of its macro-structure, as of course in its initial declaration of the theorem to be proved. Reading the detail of the proof, a reader may not have such a strong feel of it: the *lege artis* of rigorous demonstration often demand of a mathematician to work bottom-up, starting from axiom or previous necessary theorems, in a way in which many of the intermediate theorems or lemmas may appear more like porisms, i.e. quite unmotivated *vis à vis* the final goal[52].

On the contrary, in an account of a proof-as-discovered, *intentions*, *goals* and *plans* – *dreams* also, both in the metaphorical and also the literal sense[53] – predominate, giving both the story of the process of the discovery and the process of the discovery itself (the two, we saw in Section 2, have a structural similarity), a strong forward-looking drive. The proof-as-discovered appears much stronger in its cognitive logic than the proof-as-published, with constant reminders of the rationale of every step, detour, aside or parallel process, a much more holistic sense of why the prover is doing what he/or she doing *vis à vis* the goal. As soon as some kind of advanced intelligence is assumed to operate inside P-space proof, of a mathematician or, say, a proof programmer, the concept of goal must enter it: in fact, that is the main advantage of intelligence in a mathematical search, the reason why we can improve on mathematical ants as provers.

A whole conceptual dictionary can be developed regarding *goals*, concepts to describe *goal choice* and *orientation*, *primary*, *secondary* or *lesser goals*, *sub-goals*, *partial* or *approximate* goals, *contrasting* or *antithetical goals*, etc. as well as tools to measure the *achievement* of a goal, *total*, *partial* or *negative*, *plan assessment* with respect to this, *re-orientation*, *new goals*, etc. With such a toolkit, we move can unite the world where proofs occur, i.e. P-space, with proof-finders, i.e. mathematicians, using photos-and-movies as our basic language.

---

logic of discovery in their “novelistic” discursiveness than the austere, hygienic formalism of Bourbaki’s *Eléments de mathématique.* (On this, see also the next footnote.)

[52] This is much more frequent in textbooks, especially those written after Bourbaki set their high standards of rigor and axiomatic, bottom-up presentation: the theorems in their pages often appear uninvited, as by-products of a combinatorial process, rather than as necessities on the road to specific results. Thus, in Schwartz’s terminology (previous footnote), the “novelistic” elements of the older textbooks had a much stronger sense of the proof-as-discovered, which the Bourbaki people excised in the name of their Euclidean, bottom-up approach.

[53] There are two papers in our meeting with the word “dream” in their title, the first (Mazur) using it in the metaphorical, and the second (Harris) in the literal sense.

The primacy of goals in grand quests comes as no surprise. People don't, as a rule, employ random search unless they are lost, bored or desperate in a particular kind of way. Chess players don't think of "all possible responses to all possible responses to all possible responses, etc. to all possible moves" before they decide on their next move, and likewise mathematicians don't think of all possible combinations of all possible combinations of all possible combinations, etc. of symbols, to prove something. In fact, mathematicians, unlike Alice lost in Wonderland, almost always care where they get to – at least they think they do, which cognitively amounts to the same thing –, i.e. they have goals. And so, again unlike Alice, they always worry about the *way* they ought to go, forming and re-forming plans, changing them or abandoning them, creating in-between, partial and/or temporary goals to realize them.

Let us now return to our proof of:

(5) sin^2 A + cos^2 A = 1

The proof, as we read it above, is certainly convincing. But it is of the type of a proof-as-published, a bottom-up proof, advancing from the given truth (the Pythagorean Theorem) to the more advanced. As we saw, an army of randomly searching mathematical ants would have to surmount a whole combinatorial mountain, thus need lots and lots of time, and/or be amazingly lucky, before they find some kind of proof to it, to generate the sequence:

(1) → (2) → (3) → (4) → (5)

But if we begin with attempting to prove just that – i.e. if we conjecture (5), begin from there and start rewriting our conjecture in the most obvious way, thus reversing the order of propositions, we arrive much more naturally at a known truth:

(5) sin^2 A + cos^2 B = 1 →
(4) $(a/c)^2 + (b/c)^2 = 1$ →
(3) $a^2/c^2 + b^2/c^2 = 1$ →
(2) $a^2 + b^2/c^2 = 1$ →
(1) $a^2 + b^2 = c^2$

Once we've arrived at the Pythagorean theorem, we know that the new theorem, i.e. (5), is proven, we can retract our steps – this can always be done with type-1 movies, such as the one in this case. But it is knowing what we would like to end up with, (i.e. (5)), that makes the proof appear almost effortless.

4.6 Chance, necessity and structure

In Section 2, we mentioned the old Greek concept of porism (*porisma*), which the Liddell-Scott dictionary defines as "deduction from a previous demonstration, corollary, as it were a windfall or bonus". (It is pertinent to our discussion that the word comes from *poros*, meaning "passage, path, pathway", as is also the fact that this notion was brought back to the modern attention by a footnote in *Proofs and Refutations* by Lakatos, a thinker who tried to inject cognitive logic in the highly formalist tradition of mathematical philosophy.)

Of course, chance results do occur in mathematics, as a consequence of the fact that human quests do not progress by the application of rigidly programmed algorithms, and inflexible plans – as merely reading the proofs-as-written might lead us to believe – but by all the cognitively realistic methods of attempting to see, often darkly, through immense labyrinths of complexity. The concepts of intuition, inspiration, contingency, chance, error (Philby's "human factor") are all eminently applicable to mathematical research and, combined with logic and the mathematician's more rational bag of tricks, may well yield results, some of which may be indeed *windfall*.

Yet, to play for a moment devil's advocate to our own position of the importance of plans and intentions, let us turn to a mathematician as great as Sir Michael Atiyah, saying that, to him, mathematical research is anything *but* goal-oriented. In fact, Atiyah says in a fairly recent interview that: "I have never started off with a particular goal, except the goal of understanding mathematics." In that same interview he describes the way he does research as:

> "I just move around in the mathematical waters, thinking about things, being curious, interested, talking to people, stirring up ideas; *things emerge* and *I follow them up*. Or *I see something* which connects up with something else I know about, and I try to put them together and things develop. I have practically never started off with any idea of what I'm going to be doing or where it's going to go. I'm interested in mathematics; I talk, I learn, I discuss and then *interesting questions simply emerge*."

But look at the expressions in (my own) italics:

a. "*Things emerge*", i.e. problems or interesting lines of research reveal themselves: these are specific research *directions*.
b. "…And *I follow them up*". In other words, directions are as a rule *towards* something he hopes to arrive at – though it is at first unknown.
c. "*I see something*", i.e. something is identified as a goal – in fact, in the precise terminology of *"seeing"*, that we used.
d. "And I try to *put them together*…" Connecting something with something else if of course a very definite mathematical goal, a convergence of two or more type-2 arrows on a photo.
e. "…*Interesting questions simply emerge*". But questions are goals with question marks. That they "*simply emerge*" only means that they are not determined from the beginning. But they occur at some stage.

So, although Atiyah – or any other mathematician – may not "start off with a particular goal", he acquires many on the way, often in the guise of "problems". And Jean-Pierre Serre, another great, describes his research process with the phrase "I follow my nose", a goal-oriented quest metaphor if there ever was one! Whether the hunter has to decide to hunt specifically for hares or partridges *before* s/he starts out for the hunt is not that important – some do and some don't. But as the process of the chase develops s/he will eventually zero-in on the likely targets, with the right input from the environment – hares if it's hare country, and partridges if the land is friendly to patridges. But in any case, if the hunter is good, some unfortunate member of the fauna will end up on the platter.

Yet, though they do not put into question the importance of goals, Atiyah's comments, and Serre's nose, give an interesting slant to a position that may seem to suffer from too much determinism: as life is more chance-ridden and less causality-ruled than stories – remember Philby! –, so proofs-as-discovered are more chance-ridden and less inference-ruled than proofs-as-written. And this is why it is in the context of the quest that we can examine their methods much better, than in the finished, "RC" (rigorously correct) product.

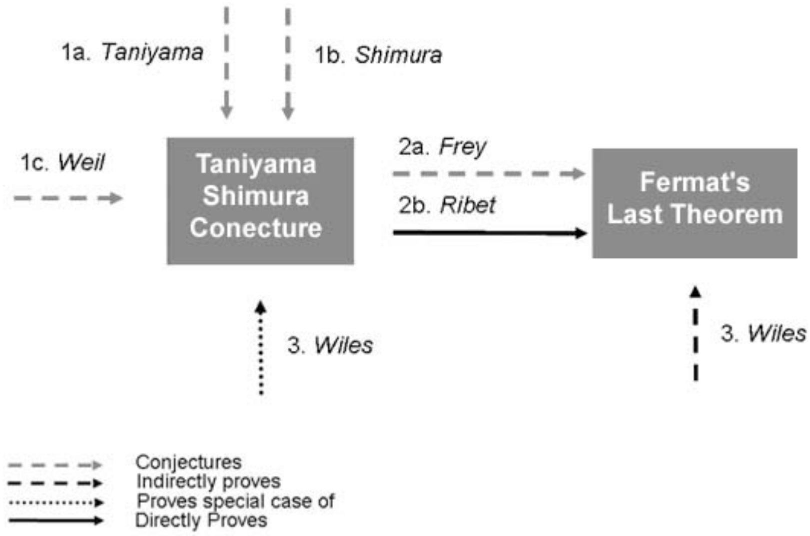
In fact, as most stories tend to overplay the causality factor, giving life more structure and meaning – except in some cases of extreme modernism specifically created to subvert determinism, i.e. the exceptions that prove the rule – so proofs-as-published overplay the strength of those magical words, "thus", "therefore", "implies", and a few more, to give more structure to P-space than it actually possesses or, more correctly, *to show in relief those* *parts of it that do possess more structure*. Both time, in the sense of history, and the collective, often Darwinian intelligence of living mathematical communities, also play their part in structuring the progress of mathematics, re-arranging what's already been discovered into much more meaningful wholes, a process whose prime example is Euclid's all-time bestseller, the *Elements*. As does, of course – though Atiyah's and Serre's opinions do not give it very much credit -- the personal goal-setting capacity of mathematicians.

But there are local *and* global structures of quests, and local *and* global ways in which they can become either more or less goal-oriented and structured. Let me give the example of the way the solution of two great problems in mathematics developed. The first is from the story of Wiles's proof of Fermat's Last Theorem. This is the full text of the abstract prefacing the paper containing it:

> *Abstract*. When Andrew John Wiles was 10 years old, he read Eric Temple Bell's *The Last Problem* and was so impressed by it that he decided that he would be the first person to prove Fermat's Last Theorem. This theorem states that there are no nonzero integers a, b, c, n with $n>2$ such that $a^n + b^n = c^n$. The object of this paper is to prove that all semistable elliptic curves over the set of rational numbers are modular. Fermat's Last Theorem follows as a corollary by virtue of previous work by Frey, Serre and Ribet.

Now, this abstract clearly creates the sense that a strong goal-orientation guided this mathematician's research plan from age 10. However, in the "gutter" between the symbols "c^n." and the word "the" lies a part of the story that had nothing to do with Wiles's personal efforts. More specifically: though proving Fermat's Last Theorem had indeed been a commendable childhood dream, when, during his University years he became aware of the unrealistic nature of such an attempt, Wiles abandoned the efforts to directly attack Fermat's Last Theorem and did not re-direct himself towards that goal until Ribet had shown that it can be attacked via the Taniyama-Shimura Conjecture, which was one of the great open problems in the field of elliptic curves. *But* – and this is a most crucial "but" – the attack on Taniyama-Shimura was only possible because meanwhile, coaxed-on by his advisor, mathematician John Coates, and for reasons unrelated to Fermat's Last Theorem, Wiles had meanwhile become an expert in the study of elliptic curves: the continuity of the "ten year old realizing a childhood dream" story is only achieved by ignoring an almost twenty-year hiatus in the goal-orientation, and a series of random events.

This is a bird's-eye view graph of the gross structure of the way the proof developed:



(The caption explains the nature of each arrow, distinguishing the first kind, i.e. *conjecturing* -- "seeing" – from the other three. The numbers of the movies refer to their chronological sequence, 1a and 1b having occurred in the mid 1950's, 1c in 1967, 2a and 2b in 1985 and 1986 respectively, and 3 started right after and completed in 1994.)

By the story told in this graph, Taniyama and Shimura and, independently, Weil, contributed to forming the conjecture, i.e. *seeing* a photo, but without being able to go there. Then another, Frey, *saw* that this, if proven, might prove Fermat's theorem, and then another, Ribet, found a path to go there, i.e. to prove Frey's conjecture that the Taniyama-Shimura conjecture would in fact prove Fermat's Last Theorem.

It was only then that Wiles, who had dreamed of proving Fermat's Last Theorem right since he was little, managed to go to (prove) a partial case of the conjecture that was strong enough for Ribet's proof to hold[54]. All this is well known, of course. What is interesting is that it is only in this, historical overview of four full decades, and the independent work of six mathematicians, that the story acquires structure. And though intention operates *locally*, in the specific arrows – thus, it was Ribet's definite goal to prove Frey's conjecture when he started to work on it (2b), and Wiles definitely wanted to prove the Taniyama-Shimura Conjecture after 1986 (3) –

[54] Though he tried to prove Shimura-Taniyama for all elliptic curves, he proved it for a special class of them, called "semi-stable", which were enough to link it up to Ribet's proof and Fermat's Last Theorem. His methods were used as a basis for a full proof of Taniyama-Shimura, now known as the Modularity Theorem, by Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor, who published the full result in 1998.

to see this diagram as representing a unified story with a central conscious hero, i.e. as the unfolding of a collective plan to prove Fermat's Last Theorem, is totally absurd, unless the hero whose intention was being worked out here was God Almighty. Or, of course, of some version of a Hegelian-style "Mathematics", an "Idea working towards its Fulfillment", or, in more mundane concepts, of the dream infesting and taking over a community of ideas. (On this, see also Barry Mazur's and Michael Harris's articles, in our conference, as well as the Shafarevitch paper to which Barry refers.) But even so, though one can make a good case for the interest in Fermat's Last Theorem (and related problems) staying alive for centuries and developing and finding a solid place in the mathematical community's collective dreams[55], the Taniyama-Shimura connection to Fermat was definitely *not* part of this dream, not until the work of people like Frey and Serre and Ribet opened that particular way.

A similar example concerns the case of the Poincaré Conjecture. Let's look for a moment at the future of the problem from the point of view of one of the most eminent mathematicians working on it in the 1970's, i.e. the great Greek topologist Christos D. Papakyriakopoulos, who devoted (at least) two decades of his life to trying to prove the Conjecture. His first attempt to reduce the Conjecture to two other conjectures, which seemed like a mighty good plan, proven (the reduction, not the other conjectures) in the early 1960s, failed because a counterexample was found to one of the two conjectures, soon after. And it is also a safe bet, that though probably hardly no other mathematician had a better feel of the intricacies and inner feel of the problem than he did in the last decades of his life, at the time of his death (1976) Papakyriakopoulos had *absolutely no idea* of how the problem would be eventually solved, 30 years later. As it turned out, this could happen only after William Thurston had formulated his Geometrization Conjecture, which Richard Hamilton and Grigori Perelman proved applying ideas partly based on the work of Gregorio Ricci, a 19th century Italian mathematician, and creator of the tensor calculus, a methodology not only foreign to Papakyriakopoulos, but also outside 3-dimensional topology, as viewed by him and the established community of advanced topologists, in his days.

So, here too one can speak of very clear goal-orientation in the proof-as-published basically operating *only locally*, i.e. within the work of individual mathematicians and/or groups for rather restricted periods of times. Thus, in historical order, Ricci, Thurston, Hamilton and Perelman – and many many others of course, not least of which is Poincaré himself – operated on their

---

[55] If the dream had not entered the collective mathematical consciousness, it's unlikely that E.T. Bell would write his *Last Problem*, i.e the book on the (then unproven) "last theorem" of Fermat, that inspired the ten-year old Wiles to make this into his life's dream. And if that hadn't happened… who knows.

own, on their local goals, and it is only in making a coarse-level historical outline that the parts fall into place and all four names can appear smoothly in the story of “the proof of the Poincaré Conjecture”: if we look at long expanses of time without a unifying plan or structure – and this as a rule imposed after the facts – predominating, it is often *the contingent that predominates*, and the ideas can be pushed into a coherent story only by unconsciously a sense of a supra-personal, mathematical ‘subject’, intent on pursuing the field’s goals, in the name of great unsolved problems or questions[56].

Finally – having finished playing devil’s advocate to myself – there are many cases where long-range goals do actually shape research for decades, or centuries. One such case is the Langlands Program, the set of conjectures put forward by Robert Langlands connecting number theory and the representation theory of Lie groups, which has created some of the most exciting research in modern mathematics, in the past few decades. And of course, there is the case of the major collective theorem of the Classification of Finite Simple Groups, a theorem proven in bits and pieces extending in hundreds of papers which collectively cover over 10,000 pages, written by at least 100 mathematicians working for over three decades. Here, the research was put in its final course by the strategic work of Daniel Gorenstein, who not only outlined the general plan of the attack for the final fifteen years of the effort, but personally played the role of chief of staff to his voluntary army of mathematicians, giving the work its central direction and spine.

So, though this does not always happen, mathematical research can on occasion have strong goal-oriented structure even in the long range.

4.7 What’s in a plan?

We are now in a position to moderate our talk of goals and plans and give it a more realistic expression. This is necessary, as I feel that one of the strongest objections to my approach may be a variation of the phrase, which some mathematicians – perhaps Atiyah and Serre among them – might uphold: “But goals and plans *are not that important* in mathematics”.

The degree and the sense in which I agree with this criticism are explained by the Fermat and Poincaré examples, just given. But the thesis about the predominance of goals becomes very strong if it is expressed in a more moderated form. *Of course* there are cases of porisms in mathematics,

[56] In Simon Singh’s *Fermat’s Last Theorem*, Andrew Wiles gives an inspiring metaphor of his years of research as a process of walking in a totally dark room, taking some months to familiarize one’s self with its layout, and then moving on to the next. In this sense too, many stages of his work were local: it is quite possible that when entering Dark Room A, he did not know that the next would be Dark Room B – this would only be apparent once Dark Room A would have been lit up by the local research.

of course an idea will often *seem* to come out of the blue – or sometimes may *indeed* come out of the blue --, *of course* some mathematicians will form their plans less consciously than others, and *of course* not all mathematicians will be like Gorenstein or Grothendieck, working out huge goal-driven projects, and others like Atiyah or Serre, following their noses or their whims[57].

But a sense of goal or plan – as a plan can very easily translated into goals, and visa versa, I treat the two practically as synonymous, for our purposes[58] –, is crucial to mathematical research at least at some level of resolution. Also, often research on a problem (which often goes hand-in-hand with general appreciation of the problem) operates at several interlocking levels simultaneously – on this, see also Section 5.2 – which is another case of the application of the Independence of Levels. But a plan, to deserve the name, need be neither totally fixed, or constant, or inflexible and unchanging. In fact, I can do nothing better in describing the way in which I mean this term than recall the words of my friend Elias Kourkounakis, chess International Master and professional chess coach, when he tried to explain to me the necessity of having a *plan* when playing a game of chess. For chess players above a certain level talk of having a *plan* as a necessity, so much so that a frequent criticism of a bad player, a *patzer*, is phrased in variants of "s/he is just moving the pieces, s/he doesn't have a plan."

This is what Elias told me, roughly reconstructed in my own words:

- A plan must serve the ultimate goal of the player in the particular game, either winning it or drawing it. (To lose, you don't need a plan!) The goal can change during the game, as for example from a win to draw, depending on the development of the game.
- A plan is a necessity, at least after a certain point in the game. (Until then, one can proceed with previously-acquired knowledge of a particular method of *opening* of the game, culled from the huge repertory of collective experience, which a strong player will, to some extent, have studied.)
- A plan can change, often more than once, during the game, indeed often it *must* – but that is absolutely no reason for it not existing in the first place.

[57] But please take also into account the earlier criticism of the last two for not being totally exact when they say they don't have goals.

[58] We can rephrase this even better as: *a plan is the outline of a goal-driven proof that has not yet happened.*

- A plan can be adaptable: sometimes the change will be in the way of modification, but sometimes a new idea may necessitate a total abandonment of the old plan and creation of a new one.
- A plan may be totally abandoned due to a new situation on the board, resulting from the opponent's moves. (This could be, in extreme cases, a blunder by the opponent, which creates a windfall opportunity for an easy win, a blunder of our own, or a totally unexpected brilliant move of the opponent which can turn our attacking strategy to a defensive one, and change the general goal, from a win to a draw.)
- A plan can have sub-plans which also should be adaptable, both in themselves and in their interrelations.

Elias's explanations of what a plan means in chess, and his insistence on the centrality of *having a plan*, despite all the qualifications, I think also answer to a large extent any question on the precise sense in which we say that "a mathematician must have a plan to prove a theorem".

In mathematics, "opponent" stands for the unfathomable complexities of P-space, and the totally unexpected things, both good and bad, it can throw up as a mathematician progresses. Of course, the mathematician does no really care if s/he loses – certainly not as much as the chess-player does – as a) no one need learn it, and b) even the direst defeat can lead to surprising insights and new paths for research – and eventual victories.

But otherwise, the analogy still stands: *of course* a mathematician may prove a theorem other than the one s/he originally started out to prove, *of course* there can be surprises, intuitions, sudden changes, *of course* s/he must be adaptable to make the most of a changing – as s/he moves – and constantly challenging environment. But these things amount to nothing else than the reality that s/he is navigating P-space, i.e. an entity so huge as to be practically unknowable from the point of view of the navigator, and thus a veritable mighty "opponent", a dark thing with – in a matter of speaking – its own will.

So, in this context too, a plan will be a set of priorities, an attempt to reduce the complexity of the quest space to manageable proportions. Starting from any one proposition, any *photo*, you can go anywhere at all, and that certainly is some *embarrass du choix*. But that's why conjectures are there, to orient research, that's why people form hypotheses, both large and small, that's why there are false starts: even in the most rational of inquiries there will always be large doses of trial-and-error. That is why, finally, many efforts to prove things fail, or change along the way.

So, plans are necessary. And plans derive from goals.

Here are a few things we can meaningfully say about mathematical goals, in no particular order:

- A goal points at a research direction, a photo *seen*, and thus reduces the magnitude of P-space into manageable proportions, i.e. creates criteria for directions which are (and which are not) relevant to its achievement.
- Research at the individual level will often start with a goal, i.e. the mathematician will think "I want to prove photo x". This goal can be changed, partly or wholly, even at an early stage, to something more or less general, closer to- or farther from the initial goal.
- Some of the goal-setting may be unconscious, i.e the result of experience (knowledge) and intuition, which is a form of unconscious thinking. (There is a big element of this in Atiyah's and Serre's comments: the greater a mathematician is, the greater as a rule is his or her intuition – and nose!)
- To reach his/her goal, the mathematician will plan a course, which will include sub-goals, i.e, will say "to prove x, I must first prove y" – these are often called "lemmas" in mathematics. (Thus, a goal is a photo, but a plan is an idea about some of the movies to be traversed on the way, and some intermediate photos.)
- The goal-setting and/or adaptation can be done at many different degrees of resolution, sometimes co-operating, sometimes not (i.e. leading to change and/or modification of goals.)
- The goal finally reached may or may not be the one actually intended, but a more or less important, or totally different one.
- The existence of goals does not mean that random search may not occur (mostly locally) at some points of the process, nor should the mathematician be unwilling to accept surprise.

Though we did not constantly refer to the photos-and-movies language, it is clear that it is particularly suitable to discuss the process and structure of proof, especially in ways that can accommodate concepts involving intentions, desires, goals and plans. The extent to which this way of looking at things, and this language, can also be useful in ways which are directly relevant to mathematical research, has not been really touched upon here – and I don't think I am the right person to elaborate on it. There are cases where graphs, or flow charts, or diagrams (which are also graphs) are used in describing or studying the structure of proofs, either in exposition, or in the context of AI, in the study of Automated Theorem Proving, Proof Planning and so on – the work of Alan Robinson is particularly interesting in this connection. Also, of course, research mathematicians can on occasion use such graphs as visual aids, thinking tools, or techniques with which to discuss their ideas.

Apart from the naturalness of this language for examining goals and goal-related concepts, the most interesting aspects highlighted by use of the photos-and-movies language in the context of proof, may lie in two related directions:

a. The issue of movies being essentially of two types, thus leading to two meanings of the word “implies”.
b. The whole question of outlines of proofs, at various levels (degrees of resolution), especially with the slant on the matter determined by the related principles of Distance, the Independence of Levels and general Self-similarity (“general” in the sense that it doesn’t imply homomorphism), as described mostly in the section of stories.

4.8 Similarities of stories and proofs, via the photos-and-movies language

Though most of what I will say here concerning the story/proof similarity is already evident to the reader – there were obvious bells ringing throughout Section 4, constantly alluding to similarities with equivalent arguments in Section 3 –, I will finish by making a general overview of the correspondences:

- Though stories and proofs are expressed in their published form via the linearity of text, they both have a strong *non-linear infrastructure*.
- This infrastructure can be expressed in the *photos-and-movies* language (a language of graphs) for both stories and proofs. In both worlds, photos are about stasis; movies are about movement – which occurs in time.
- The general setting for stories and proofs is *E-space* and *P-space*, respectively. Both are infinite graphs in which all potential stories and proofs can develop. Both look very much alike, if we forget what their respective photos and movies represent.
- The totality of these humongous spaces is never dealt-with directly. We only deal at any one time with more manageable subgraphs.
- Both stories and proofs can be thought of as *dags* (see p. 40), possibly with some loops. in E-space and P-space, respectively, ending in the conclusion (root of the graph) marked with the equivalent of “The End” in a story and QED in a proof.
- The *linear form* of a story or proof is in essence a *linearization* of the dag (possibly plus some loops) of the story or proof. In linearizing, we

lose information, yet linearize we must, due to the structure and limitations of our cognition which only handles linear symbolic forms effectively.

- Storytellers create stories, mathematicians create proofs as linear artifacts, most of whose information is contained in the concatenation of photos. But the photos are not as important as the graphs, i.e. photos and movies combined, underlying them. Stories are never just about the events, nor proofs about the propositions, but about the way we traverse them, and can use these to understand ("*go to*") others.
- In creating stories and proofs, storytellers and mathematicians are endlessly familiarizing themselves with E- or P-space, learning to know it better by setting destinations and finding their way in the immensity.
- In this sense, beyond having a *fabula* and *syuzhet*, a story also has an *Ur-fabula*, the non-linear structure in which it is lodged. These three terms, Ur-fabula, fabula and syuzhet can also be directly applied, via the photos-and-movies language, to proofs.
- Both stories and proofs use *wishes*, *desires*, *intentions*, *goals* and *plans* to deal with the immensity of E- and P-space. Though searching for ways to *go to* a photo is crucial in both processes, stories and proofs, goal-setting also requires *seeing*.
- There are two kinds of movies in both E- and P-space, which we can call *type-1* and *type-2* (it was for these two types, in the context of proofs, that we used the symbols ☺ and 😐, respectively.)
- We can also apply a concept of *weight* to movies. Type-1 movies have weight equal to 1 in both worlds and are *strictly transitive*. Also, stories or proofs consisting solely of type-1 movies are essentially *paths*, directly linearizable, without loss of information.
- The weights of type-2 range between 1 and 0, while stories or proofs containing type-2 movies are trees. The transitivity of type-2 movies is weighted, the effect (weight) of a path of type-2 movies being equal to the product of the weights.
- We cannot create *optimal length linearizations*[59] of type-2 subgraphs without significant loss of information. A corollary of this is that:
- Contrary to a subgraph consisting strictly of type-1 movies, a subgraph containing type-2 movies cannot be linearized in a unique way.
- The multiplicity of *outgoing movies* from photos has to do with *choice*, i.e. the fact that we can go from one photo in a story or proof to many others, a result of their being lodged in the richness of E- or P-space.

---

[59] *Optimal length* in this case means that *the linearization of a path does not contain too many more photos* (in the sense of multiple occurrences) *than the subgraph itself*.

- The multiplicity of *incoming movies* on a photo in a story or proof graph has to do with the nature of type-2 movies, in both cases.
- *Outlines* are an important aspect of both stories and proofs. In photos-and-movies language, these can be *excisions*, *contractions* and *combination* outlines. Outlines can operate at many levels. Thus, assuming a particular version of a story as a root, and then creating gradually coarser levels of outlines we can form a tree of outlines, having many branches at every level, and many levels.

  Using outlines, we can talk of three important characteristics of stories and proofs.
- A) Both conform to a *Distance Principle*: the closer two photos are in the linear form, the likelier it is that there exist a high-weight type-2 movie, or high weight path of type-2 movies, connecting them – this is a statistical law, which has exceptions. If two far away photos are connected through type-2 movies, this movie also appears in outlines. The higher the weight, the higher the level of outlines in which it may appear.
- B) Both operate via some form of *Independence of Levels*: the local structure of the graph is usually partly independent of the global (outlined) structure. This is also a statistical principle: there are exceptions, and when they are they are important – and survive in outlines.
- C) Both are in some way structurally *self-similar*: if you look at the graph of an outline of a story or proof, without knowing that it is an outline, you could mistake it for a graph of a story or proof itself.
- Both in E- and P-space, we can speak of internal transformations and equivalences. Thus, stories can look like other stories – and this means that their graphs are at least partly homomorphic – and so can proofs. Thus, the concepts of *variation*, *adaptation*, *genre*, *patterns*, *motifs* (in the story sense), and so on, exist in both spaces. (In mathematics, the concepts of category theory, and especially of functors, are particularly suited to describing many of these relations.)
- For both stories and proofs, we can speak of the story-as-created and proof-as-discovered, in the first case, as opposed to the story-as-written (or story-as-published) and the proof-as-published. (The concept of successive *drafts* is also relevant here, reflecting in-between stages.)

There are obviously also many differences, mainly having to do with the subject matter of stories and proofs, as well as their purpose, which are as often as not related both to the psychology and sociology of their knowledge, i.e. the way that the individuals and social groups by which they are developed and used are constructed and in which they handle stories or proofs. But as we are here speaking of a level that is closer to the formal and structural, the main difference between stories and proofs – in this sense – is the one reducible to the employment, for stories, of the much richer, more ambiguous and thus much more flexible instrument of natural language, while proofs are created through the much more restricted, and much less ambiguous tools of the technical languages of mathematics.

* * * * *

The similarities noted, whose definition and description relies heavily on the photos-and-movies language, are too many and too strong to be accidental or insignificant, for them to be just mirages created by a procrustean formalism.

In addition, a strong intuitive argument for the similarity of stories and proofs is the ability that human beings possess, in both cases, for living cognitively inside huge graphs, whether E-space or P-space, which one cannot know except partly and linearly – because of our cognitive make-up, we can only process representational symbols linearly, either to interpret ("read") or to produce ("write") – yet adequately enough to discover meaning, and exist inside of without a sense of disorientation. Storytellers and/or users, and mathematicians inhabit these huge non-linear humongous spaces naturally, possessing non-linear knowledge (i.e. a sense of "place") which they augment, with time, through a constant increase their experience of paths, i.e. stories and proofs. These, they are constantly translating into some kind of deeper knowledge of the non-linear structure, first trees, then dags, and then, ultimately, highly non-linear knowledge of the graph structure. The linear experience, perception and practice constantly interacts with, and augments, a deeper non-linear knowledge – and visa versa.

But my strongest argument for the significance of the similarity of stories and proofs goes *outside* the photos-and-movies language: in fact, it is the very reasoning by which this language is constructed in each case, for stories or proofs. This is that the action sentence and the basic mathematical proposition seem to be quite naturally, i.e. in a cognitively meaningful way, the basic, irreducible entities both of the transduction of reality into narrative and the fixing of it in manageable formal rules. Once this is accepted, the operations on them via the invisible movies yield processes that are so strikingly similar, both locally and more globally inside the huge graphs of stories and proofs.

This seems to me to be the best evidence of a similarity which indicates a family relationship.

5. Two stories in photos and movies

This section contains a trace of two genealogical arguments, the one evolving in historical, the other in pre-historical time, that should be significantly longer to begin to have a hope of being convincing. Yet, having exceeded, by far, a reasonable length for this paper, I will give here but mere indications.

5.1 The old story

Though Babylonians and Egyptians used had developed quite complex computational mathematics, the logico-deductive method which forms the signature style of mathematics, including the basic methodology of proof and the axiomatic construction of theories, is almost exclusively Greek. If we take Euclid's *Elements* as the expression of this process in its maturity, and go back to the earliest extant evidence of proofs, we can demarcate the process of this transformation from the earlier to the Greek style as contained quite comfortably in a two-hundred year period, from sometime in the first half of the fifth century BCE to the first decades of the third – and it's very possible that, with more evidence, the limits can be narrowed down to a period half that long.

The process of this transformation is very fast, and extraordinarily successful by any standards, especially so if we consider the fact that the general form of mathematical demonstration and theory construction as seen in Euclid has not fundamentally changed very much since then. Or, to put it less controversially: though this form has been significantly amended and improved upon since then, the basic methodology invented by the Greek mathematicians is still there. And, equally importantly, practically *all* its proven truths are still considered to be true. (To understand how impressive this is we must remember that this cannot be said to be even remotely true of any other field of ancient knowledge.)

Of course, a lot of this persistence in time of mathematical truth is due to the nature of mathematical epistemology. But as this epistemology was developed precisely in this period, this argument strengthens, rather than diminishes, the enormity of the achievement. The awe at the extreme success of this genesis becomes even greater if we consider the "big bang scenario" (this is my terminology) put forward by some prominent scholars in the field and recently well outlined and defended by one of its strongest proponents, Reviel Netz, in his important book, *The Shaping of Deduction in Greek Mathematics*. According to this scenario, there is a more or less sudden birth of the new paradigm, rather than a slow, gradual development from the older one. The new mathematics is suddenly born, somewhat like Athena jumping out fully-formed from the head of her father, Zeus: this probably occurs over a few decades, or even years, most probably sometime late in the fifth century[60].

I want to propose the hypothesis that the reason why the new entity could come into existence so suddenly, and acquire such extraordinarily well-developed form in the span of a few decades, was that it was more or less directly modeled, cognitively – though at least partly *unconsciously* so – on the pre-existing, highly developed discipline of storytelling, which had been natural to humans for at least some decades of millennia before the time of classical Athens, but which had also reached great heights in Greece in the previous for centuries. This was a discipline for which there was obviously a strong cognitive infrastructure already in place.

From our knowledge of the facts describing the context of the birth of Greek mathematics, I highlight some significant elements, as background to this hypothesis:

- The greatest part of the transformation from the old to the new mathematics most probably occurred in Athens, most probably starting sometime in the 5th century.
- The tendency towards critical thinking, without appeal to higher, extra-rational authority, had developed particularly strongly in Athens during the 5th century, largely as a result of the new political and social organization of the *polis* and, more especially, of democratic government (on this, see especially the work of Jean-Pierre Vernant and G.E.R. Lloyd)
- This tendency was particularly enhanced by the dual processes of political debate and forensic argument, exercised in a spirit of liberty.
- This tendency left strong marks on other facets of cultural life, before affecting mathematics, especially literature (see the literature on the influence on forensic argument style on classical tragedy), and also the related discipline of rhetoric; this last influence was also an important factor for the development of philosophy, especially in its departure from some of its poetic, pre-Socratic expressions, towards the more rational, early Sophistic and Socratic/Platonic form.

The Hungarian scholar Árpád Szabo, in his book, *Anfänge der Griechischen Mathematik* (1969), proposes a philosophical origin for the deductive method, tracing its origins in the Eleatic School, and more specifically the writings of

[60] In other words, around the time Plato was born. In fact, Reviel Netz describes the period of this change as coinciding more or less with the lifetime of Plato (429-347 BCE).

Parmenides in which we do, indeed, meet the first instance of indirect proof, and an appeal to universals. But, despite all the interesting and valid insights contained in Szabo's book, G.E.R. Lloyd points out that to wonder whether Greek mathematics could have begun in philosophy is rather pointless: we must remember that we are talking of a time when neither "philosophy" nor "mathematics" were strictly defined and/or delimited and as a rule the practitioners of the one were also – at the very least – well-versed in the other. (The *ageōmetrētos mēdeis eisitō* --"no one ignorant of geometry shall enter– over the entrance of Plato's Academy is a good case in point.)

If we look at the picture of intellectual practices in 5th century Athens given by Vernant and Lloyd, as well as Socrates' vilifying of the Sophists – its obsessive intensity is as good an indication as any of their influence –, it seems that the main burden of the passage from less to more rational thinking during the 5th century, culminating in the creation of the strongly logical system of thinking whose prototype is mathematical proof, was mostly carried out in, or around, the practice of public speaking. Oratory and dialogue were the main tools of persuasion, i.e. proof, in both law and politics, which later solidified into the accepted disciplines of *rhētorikē* and *dialektikē,* both of which rely on more or less the same basic toolbox. But even before the elucidation of clear-cut theories of these practices, there was the development of a sophisticated practice[61]. And I think this is to a large extent the vehicle – or, if not the vehicle, then at least the reflection – of the transition from *poiētikē* to *philosophia*, i.e. from the mythic (poetic and narrative) mode of thought to the rational (including both philosophy and mathematics), whose *palaia diaphora* ("old difference" or "quarrel") Socrates alludes to. The intensity of his dislike of the narrative mode, in the context of this discussion in the *Republic*, makes us think that the abandoning of stories for the benefit of arguments is almost the origin myth of philosophy, a form of parricide, a symbolic killing of the old paradigm – this is the culmination of the "old quarrel" – which sets the new one free.

Obviously, storytelling existed long before the time of classical Athens in the practice of its ubiquitous, quotidian form, and we have many pre-5th century good literary Greek stories. We also know that proofs appear sometime in the late 5th and early 4th century. Of course, Greek proofs at this time are mostly geometrical, and a lot of the characteristics of this new mode of persuasion can be attributed both to the intuitive obviousness of the basic concepts regarding its subject matter, i.e. space, the reinforcement by of this intuitive obviousness by the tool of diagrams ("to see is to believe"; "a picture is worth a thousand words") as well as to the social practices of new

[61] The attribution of the first manual of rhetoric to Corax and Tisias, in early 5th century Syracuse, has recently been contested by some eminent scholars, who argue for a much slower, longer and less theoretically self-conscious beginning.

communities of high-caliber intellectuals, with their high standards of precision in definition, as well as confirmation and verification, particular to mathematics[62]. But let me very briefly outline here, using the photos-and-movies language, a short narrative of how the insights of Sections 3 and 4 can play a big part in helping us understand better the transition from storytelling to the advanced deductive constructs characteristic of Greek mathematics as evidenced, say, at the time of Euclid.

Let's first look at public speaking in Homer. The speeches of the Homeric heroes, written and/or recorded from oral sources approximately about two hundred years before the birth of proof, are essentially proto-rhetorical. There is an essential difference between these and the sophisticated discourses of the Athenian orators in the 5th and 4th centuries: of the three elements, *pathos*, *ethos* and *logos*, that were later described as the basic dimensions of oratory, the third only appears in Homer in a rather primitive form. Interestingly, what there is a lot of in the earlier speeches, apart from the three elements of *ethos*, *pathos* and *logos* – in fact, there is much more of it, in proportion, than in the rhetorical speeches of the 5th century –, is narrative-type speech, both in the forms of sentences describing basic actions, and in longer narrative sequences, mini or bigger stories.

Of the three recognized rhetorical elements, *ethos* is prominently there, most of it in the form of self-aggrandizement. *Pathos* is dominant (look at the quarrels between Achilles and Agamemnon and Achilles and Thersites in Book One of the *Iliad*, as examples), a lot of it in the way of *ad hominem* arguments[63], as well as many appeals to the emotions, both of the speakers and the listeners. There is of course already evident in the Homeric speeches the very characteristic of later discourse energy of trying to *enforce one's will through words,* set in a discourse among more-or-less equals. In fact, the Achilles-Thersites discourse in Book One of the *Iliad* is an exception that proves the rule: that it ends not in logical victory but in a good beating is due to the fact that the argument occurs between representatives of different social classes, and thus, when the socially superior (Achilles) is tired of arguing he can physically abuse the inferior (Thersites), to everyone's, except Thersites', delight.

But there is little *logos* in Homeric speeches, at least in the sense of structured demonstration, and what there is of it is far away from the sophisticated, elegant intellectual constructions of the 5th century, not just the elaborate rhetorical conventions that we meet already in Gorgias – i.e., those

---

[62] On these matters, see the work of Vernant, Lloyd, Netz and Szabo, as well as the fiery criticism of Sabetai Unguru, "On the need to rewrite the history of Greek mathematics". Also, the chapter on ancient Greek mathematics and philosophical schools in Randal Collins' unjustly neglected, in the mathematical historical circles, *The Sociology of Philosophies*.

[63] My personal favorite is Achilles calling the elder son of Atreus *oinobares kunos ommat' echôn*, *kardiên d' elaphoio* ("you drunkard, with the eye of a dog and the heart of a deer").

whose likes Aristophanes had such a great time lampooning, more especially in the *Clouds* –, but also the more basic forms of the syllogisms.

Significantly, what there is of *logos*, or more logically-structured argumentation, is in the way of someone appealing to a more general principle, often divine decree, parental admonishment or proverbial wisdom, and urging the hero to act according to it. Thus, the sentences which I've underlined in Peleus' – Achilles' father's – words, which Odysseus quotes in the scene where he tries to persuade Achilles to return to the fighting, include a sort of major premise, which the ultra-cunning speaker applies to the situation at hand (*Iliad* IX, 311-322; translated by Ian Johnston.):

> My friend, that day your father, Peleus,
> sent you off, away from Phthia,
> to join Agamemnon, didn't he say this:
> "My son, Athena and Hera will give you
> power, if they so wish, but *you must check*
> *that overbearing spirit in your chest.*
> *It's better to show good will, to give up*
> *malicious quarrelling*. Then Achaeans,
> young and old, will respect you all the more"?
> That's what your old father said, advice
> which you've forgotten. So even now
> you should stop, cease this heart-corroding rage.

But if we look at speeches in 5th century literary texts antedating the earliest we have of rhetorical orations, such as those occurring in the works of Aeschylus or Sophocles, or in Pericles' Funeral Oration, recorded by Thucydides in his *History* and given in 431 BCE (i.e. seventeen years before our earliest extant purely rhetorical text, *Gorgias' Encomium of Helen*), we see that the situation has changed. Look for example at the scene where the same Odysseus is trying to convince Neoptolemus to play along with him, and con the ailing Philoctetes out of his all-powerful bow, in Sophocles' *Philoctetes*. The argument is long and elaborate, with Neoptolemus, expertly, though subtly led-on by Odysseus' cunning, maneuvered out of his initial, strong moral stance of refusal, to a state of more or less resigned acceptance. In these, we see three elements of a methodology that we could almost call pre-Euclidean in three of its dimensions: a) *dealing with definitions*, as in the subtle variations between 'lie' (*pseudos*) and 'deceit' (*dolos*); b) stating the appropriate *first principles* as in this exchange (my underlining, in translation by Robert Torrance):

NEOPTOLEMUS: Do you not think that *telling lies is shameful*?
ODYSSEUS: No - *not, at least, if lies lead on to safety*.

Or, finally, in c) the elaborate structure of a long argument, using many intermediate deductions, necessary to lead, step by step, to the final result.

Look at the simple narrative sentence made famous by E. M. Forster in his *Aspects of the Novel*, as an example of what he calls *plot*:

The Queen died and then the King died of grief.

This is a mini story, pure narrative, a strong *c*-movie connecting two photos. Let us now adapt it to a more democratic form, also reversing the order in which the two spouses die and giving it some historical detail.

(1) Chrysippus was killed in war, and Hermione his mother died of grief.

Again, a mini story. Sentences of this form are often used in rhetoric, often brought in as instances of the accepted convention whose name is *paradigma* (example), when argument is made by analogy with a specific, known case. While it is natural that (1) could be straight out of the third- or first person narrative in a story, the second person (2) could only be fit in a dialogue, where one character is trying to persuade the other not to go to war – or, of course, in a rhetorical speech addressed to the orator's audience:

(2) Even if you don't care about your own good, do you want to go to war, get killed, and have your mother die of grief, like poor Hermione when Chrysippus her son died?

Now, let's focus on this sentence and extract from it what in rhetorical parlance is called an *enthymema* or *enthymeme*[64]:

(3) If you get killed in war, your mother will die of grief.

The transition from the italicized part of (2) to (3) is basically the transition from a question to a declarative sentence, a simple play with the grammar. But the huge, though not so visible, change has occurred already in the

[64] A sequence of truncated syllogisms is what's called in historical parlance a *sorites*, a name that would apply quite well to the Gavrilo Princip argument in Section 3.

transition from (1) to the italicized part of (2), which we can also happen as a direct transition from (1) to (3), omitting (2). Let's rewrite (1) as (1*) and (3) as (3*), making the sentences pure photos-and-movies by substituting with an arrow the linguistic markers of the movies, i.e. the words "and" in (1) and "if" and an (implied) "then" in (3):

(1*) Chrysippus is killed. $\longrightarrow$ His mother dies of grief.

(3*) You will be killed. $\longrightarrow$ Your mother will die of grief.

This rewriting shows clearly their structural similarity. But there is a big difference, though quite subtle, because it is yet half-formed, the trace of which we can see in the linguistic markers: the movie in (1*) is a *c*-movie, i.e. it indicates pure causality. But in (3*), by virtue of the construction being hypothetical, *there is already an element of implication*. To see this better, substitute the movie in each case by "and thus": the substitution makes (1*) look a lot less like (1), than (3*) looking like (3).

Yet, though it is clear that (1) is a purely narrative construction, it is not so clear that (3), the *enthymeme*, is purely logical. Look at some definitions of the *enthymeme*: "a syllogism in which one of the premises is implicit" (*Merriam-Webster Dictionary*); "the informal method of reasoning typical of rhetorical discourse. The enthymeme is sometimes defined as a 'truncated syllogism' since either the major or minor premise found in that more formal method of reasoning is left implied." (*Silvae Rhetoricae* at *http://rhetoric.byu.edu*); "a deductive argument, especially a categorical syllogism, from whose ordinary-language expression one or more propositions have been omitted or left unstated." (*A Dictionary of Philosophical Terms and Names*); "an abridged syllogism, one of the terms being omitted as understood." (Richard A. Lanham's *Handlist of Rhetorical Terms*).

It is interesting to note here that all of the definitions involve an *omission* -- best expressed in the phrase "truncated syllogism" –, as if the enthymeme historically comes *after* the syllogism. But this is not reasonable, as the syllogism is obviously the more advanced form, both logically and, one would think, historically: an enthymeme is really narrative, action sentences connected with causal links (though perhaps in the future tense, and/or hypthetical mode), that could only be said to be veering towards the deductive because we view them as "truncated syllogisms". But a syllogism is a much more formal construction, possibly also developed after a strong influence of the practices of writing, as mentioned also in Netz's book. And in fact we do meet enthymemes long before syllogisms, in the extant literature of ancient Greece – and eminently so in Homer.

Walter J. Ong's works on the oral vs. the written, as well as his discussion of rhetoric, give us important insights on these interactions. I quote more particularly from *Rhetoric, Romance and Technology*, after first reminding the reader that *enthymeme* comes from *en*, the preposition signifying "in" or "inside", and *thymos*, the word for "soul". Here is Ong, talking of the enthymeme:

> It is thought of as concluding because of something unexpressed, unarticulated: *enthymema* primarily signifies something within one's soul, mind, heart, feelings, hence something not uttered or "outed" and to this extent not a fully conscious argument, legitimate though it may be. Aristotle's term here thus clearly acknowledges the operation of something at least very like what we today would call a subconscious element.

But this subtle pressure of the hidden, the subconscious, whether repressed, merely unspoken or purposely ignored, is very frequent in language. In fact, language often hides a lot of its premises, or rather – to do away with the intention in the verb "hides" – it *leaves many things unspoken*. Yet this is exactly the loss of information we described as happening in the process of linearization, going from a complex non-linear web of movies to a linear and much smaller (in terms of information) form, as expressed in the words of the story itself.

In fact, let us see what is the logical information lost in our example of the *enthymeme*. But no: rather than accepting the enthymeme as subsequent to the syllogism, let us go the other way, and see how *we make the enthymeme into a syllogism*. Here it is, with the dotted movie now coming from the previously unspoken photo, with all three photos named according to their function in a logical syllogism. (That the major premise is not an inductively verifiable truth does not here affect the logic of the construction – though this would make a *perfect point of attack*, if one wanted to attack the veracity of the conclusion.)

(5)

<u>Major premise</u>
Mothers die of
grief when their
sons are
killed in war

<u>Conclusion</u>
Your mother will die of grief.

<u>Minor premise</u>
You will be killed
In war.

But if rather than think of the enthymeme as a syllogism from which the major premise "has been omitted", we assume the growth of conceptual sophistication also to be reflected in genealogy, we get this historical sequence:

Narrative
("Chrysippus was killed, his mother died of grief.")

↓

Paradigma
("Would you like it if your mother died, because you were killed in war?")

↓

Enthymeme
("If you are killed in war, your mother will die.")

↓

Syllogism

It is only the last jump that occurs also within the province of a self-conscious art of rhetoric, in 5th century Athens, created, developed and fine-honed in practice – before theorists such as Aristotle codified and described it –, in the political bodies of the *boulē* and the *ecclēsia* and the judicial institution of the *Heliaia*, where the practical and specific has to interact with the general, as expressed in the city's laws and/or the accepted cosmology. This kind of derivation of the particular from the general is beautifully seen as a metaphor of the way law functions in the *polis* in Heraclitean fragment DK 114: "Speaking with reason (*logos*) one must be strengthened by that which is common to all, as a city is by its laws, and even more so." (*xun noō legontas ishurizesthai hrē tō xunō pantōn, okōsper nomō polis kai polu ishuroterōs*; *ξὺν νόῳ λέγοντας ἰσχυρίζεσθαι χρὴ τῷ ξυνῷ πάντων, ὅκωσπερ νόμῳ πόλις, καὶ πολὺ ἰσχυροτέρως.*) And it is this interaction between the "common to all" (*to xunon*) and the views of those who want to live as if they "have their own mind" (*idian ehontes fronēsin*)[65] that is constantly debated in tragedy,

[65] Reference to the Heraclitean fragment DK 2, which readers will have met prefacing T.S. Eliot's *Four Quartets*: "Although reason (*logos*) is common to

especially Aeschylean and even more so Sophoclean, mostly in the interaction of the Chorus, representing *to xunon*, and the protagonists, who have grand doses of their own mind.

In the works of Aeschylus and Sophocles, and especially in the interaction of the chorus and the protagonist – less so in those of Euripides, where the function of the chorus has partly changed –, or in the *agon* between characters with contrasting opinions, as for example Antigone and Creon, there is very frequent appeal to general principles, either in the form of divine law, proverbial, unwritten wisdom or the laws of the polis. Though the interrelation of the particular to the general is not yet codified into the basic syllogism, we have a strong sense of the interaction from which it sprang.

It is with a view to this interaction that the transition from the narrative to the logical mode occurs, precisely by reversing the normal process of narrative and going back, from the *linear* to the *non-linear*, i.e. stating the obvious-seeming general principle (the major premise) that would have been lost in linearization, i.e. the one that makes the enthymeme into a syllogism, *by making the unconscious, conscious*.[66] So, when a rhetor's *paradigma* or *enthymema* is under attack, in free discussion, it is the strength of the implication from the combined *major* AND (Boolean) *minor* premise which makes it unassailable – as long of course as the major premise *is* generally acceptable. In fact, this is perhaps the most basic reason why mathematical thinking reaches such heights of certainty, compared to any other discourse in classical Greece: that because of its clear focus on (also diagrammatically described spatial phenomena) its premises, as we later see them in Euclid's *aitēmata* and axioms, can be generally acceptable as "obviously true".

The structure of the syllogism as exemplified in diagram (5), on page 94, contains two type-2 movies, as described in Section 4. It is not a structure of type-1 movies, where logical derivation needs but one incoming movie, but the more complicated form of *thus* or *therefore*, which needs *at least* two[67].

What makes this kind of discussion even more interesting is noticing that one would be hard-pressed to decide whether the dotted movie in syllogism (5) is implication or causation – or, for that matter, where there is not an element of both in the lower one as well. What emerges from this is that *the language of photos-and-movies is perhaps stronger than the content*

all, the many (*hoi polloi*) live as if they have their own mind." (*Tou logou de eontos xunou, zōousin oi polloi os idian exontes fronēsin*; τοῦ λόγου δ' ἐόντος ξυνοῦ ζώουσιν οἱ πολλοὶ ὡς ἰδίαν ἔχοντες φρόνησιν.)

[66] In this context, "unconscious" refers more often than not to what cognitive scientists call the "cognitive unconscious", rather than to a more psychoanalytically-flavored sense of repression.

[67] Netz rightly stresses the prevalence of the word *toutestin* ("therefore") in Greek mathematical texts.

or meaning of the movies themselves. In fact, one reason for the strength of the photos-and-movies language is that we do not have to refer to such content: rather than appealing to any kind of essence of the temporally-developing process of the movie, we can base most of our discussion on the purely formal consideration of the existence of a V-shaped, rather than an I-shaped, linear junction.

This is seen even more, if we further break down the major premise adding another movie. In this case, the major premise appears to be just the simple declarative sentence of the form of (1*), an advanced *paradigma* acquiring extra force by induction, i.e. by the larger number of cases of it occurring giving it some kind of greater statistical validity.

(6) <u>Major premise</u>

Sons die → Mothers die of grief

<u>Conclusion</u>
Your mother will die of grief.

<u>Minor premise</u>
You will be killed.

We now have a mini graph, with a V-junction, all three movies of which are more causal than deductive. Yet, this is the quintessential logical argument, as developed in classical Greece, a structural equivalent of the all-powerful *modus ponens*: and this is shown here to be nothing more than bringing out the usually unseen non-linear structure in which a narrative sentence is embedded, i.e. a subgraph of E-space to which it belongs, and which contains the relevant relationships to reveal its exact causal structure.

This is a process of going from the relaxed, informal pace of storytelling, and the richness of natural language, to the more exact rendering of bringing also to the light the unspoken, and exhibiting the full non-linear structure of the underlying graph for inspection and control. Seen in the context of individual psychology of the person going for such a style, we could brand it as *neurotic*, or *obsessive-compulsive*, or, alternatively – and perhaps synonymously – an almost *ritualistic* process of saying all the necessary things at every stage. And there is certainly an element of that in mathematical explanation, compared to other forms of discourse. In fact this pedantry (to an outsider) of always being exact, and insisting on stating the obvious and/or unnecessary, is to the layperson the trademark of mathematical style, that which provides the material for most of the jokes about the absent-minded, nerdy mathematician.

I find particularly interesting in this context, Netz's comment, in his *Shaping of deduction*, about finding Euclid excruciatingly boring. Yet, the boring or pedantic style in the *Elements* – which many mathematical readers also see, to greater or lesser degrees, in such works as the Sisyphean *Principia Mathematica* of Whitehead-Russell, or the hygienically austere *Eléments de mathématique* of Bourbaki –, is precisely the side-effect of the urge to convince absolutely of the truth of one's position, once the elements of *ethos* and *pathos* (let alone beatings!) are not allowed the prominence they still had in the archaic world of Homer. Such a motive makes for over-explanation and over-expansion leading to a style that can be justifiably thought of as neurotic – and nothing is as boring as (other people's) neuroses.

In the context of late-archaic and classical Greek culture, the flowering of this style is mostly a reaction to the social pressures met in a democratic society in both political and forensic discourse, where free citizens have to convince equals with argument, rather than rely on force or appeal to the authoritarian whims of a despot, human or divine. But whether psychologically or socially determined, it is from this bringing-to-the-light of the unsaid, of speaking the unspoken, i.e. of the non-linearization of the linear, mostly in the form of the general rules (i.e. in the context of specific arguments, the "major premise"), that creates the first logical-deductive tools.

If we look at an early sample of a speech that is self-consciously rhetorical in this era, like say Gorgias' *Encomium of Helen* (ca. 414 BCE), the constant appeal to general principles has acquired an almost ritualistic flavor. By now, the application of formal rules has become codified (*ritualized?*) into a repeatable form, where the need of convincing others has developed into formal habit – of course, this element also has to do, at least in the particular case of the *Encomium*, with the influence of the style of an expressly *epideictic* speech, a sort of show piece for a master rhetor's art, i.e. a speech unconstrained by the pressures of a real-life political situation. But the most important element to notice is the prevalence of general arguments, i.e. those that would appear in a syllogism, as a rule, as the major premises. This barrage of generalities is what would mostly distinguish this style of theoretical discourse – a style that is, if Socrates will excuse us, closer to *philosophia* than *poiētikē* – from earlier, pure narrative discourse.

This is how the *Encomium* begins (transl. Brian Donovan):

> The order proper to a city is being well-manned; to a body, beauty; to a soul, wisdom; to a deed, excellence; and to a discourse, truth--and the opposites of these are disorder. And the praiseworthy man and woman and discourse and work and city-state and deed one must honor with

> praise, while one must assign blame to the unworthy--for it is equal error and ignorance to blame the praiseworthy and to praise the blameworthy.

But let us now look at a speech by an active politician, the famous Funeral Oration of Pericles, given in 431 BCE, and recorded by Thucydides in his *Peloponnesian War*. Though rules and/or other types of rhetorical rituals (as those applying specifically to the genre of Funeral Oration) may apply in it, the "boring" logical style of battering the audience with your premises, is much less prominent. The kind of arguments that predominate in its surface form are the "truncated syllogisms" of the enthymeme.

Let's look, in particular, at this bit of it, in a translation by Richard Hooker:

> I have dwelt upon the greatness of Athens because I want to show you that we are contending for a higher prize than those who enjoy none of these privileges, and to establish by manifest proof [68] the merit of these men whom I am now commemorating. Their loftiest praise has been already spoken. For in magnifying the city I have magnified them, and men like them whose virtues made her glorious. And of how few Hellenes can it be said as of them, that their deeds when weighed in the balance have been found equal to their fame! I believe that a death such as theirs has been the true measure of a man's worth… For even those who come short in other ways may justly plead the valor with which they have fought for their country; they have blotted out the evil with the good.

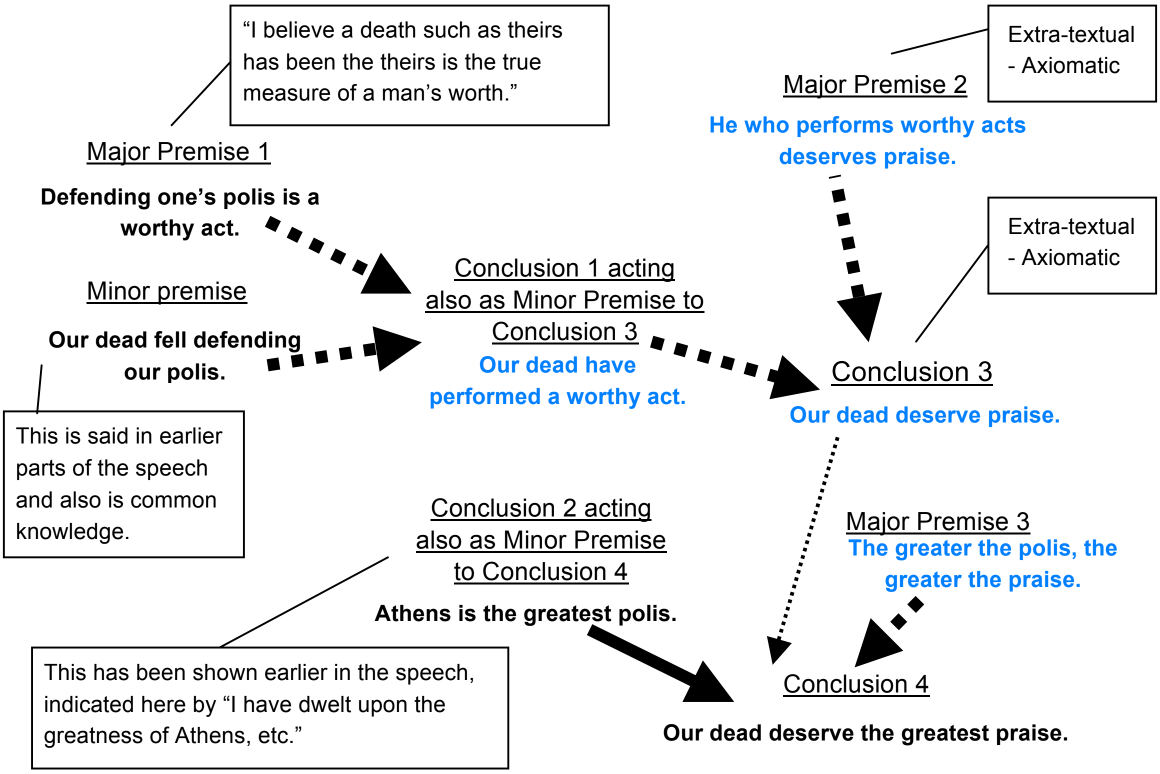
Now, Pericles was no doubt a very capable orator and, what's more important, one schooled, through experience if not also instruction, in the best techniques that Athenian political discourse had to offer. This part of his speech could definitely not qualify as pure narrative, though it does not have the turgidity of the professional sophists' demonstrations high argumentative style. What is so interesting here is that Pericles in essence presents here – to rephrase this in the photos-and-movies language – a *linearization of a much more complex non-linear argument*. And it is safe to assume that both he, and his audience, which on this occasion was probably a very substantial part of the citizenry of Athens, could be familiar with this structure, i.e. that it was both within his capacity to create, through experience, this linearization,

---

[68] As the word "proof", which most translators use at this point, is unusually loaded in a discussion which also concerns mathematics, we must make it clear that it is not the word usually used for mathematical proof. The phrase translated here as "to establish by manifest proof" in the Greek is *phaneran sēmeiois kathistas*, which translates, in a more literal sense, as "make manifest through signs", or "show forth (demonstrate) through signs". (I cannot attest to whether any of these words is also usually employed in mathematical texts.)

and for his audience to consider his point as demonstrated (or: manifest), a demonstration that can only be taken to be true if the whole, non-linear structure, containing the syllogisms of which we see the truncated forms in the speech, is somehow available to cognition. Clearly, both the passage from linearization to non-linear interpretation were mostly *unconsciously* available to the listeners and probably partly consciously to the speaker.

This is a very rough sketch of the non-linear structure underlying the above part of the Funeral Oration:



Here, the premises printed in black, some of which are major and some minor, are those actually appearing in the Oration, in which their linear sequence (in the surface form of the speech) does not always reflect their logical sequence; but those in blue show the parts of the syllogisms that *are not stated in the surface form*. The boxed texts provide their justification, which are either in the text or, in the cases of Major Premise 2 and Conclusion 3, are taken as axiomatic, i.e. "things too obvious to prove".

Our central point is that, regarding the above from the point of view of logical argument, the final Conclusion 4 *cannot be arrived-at without the* (blue) *unstated premises and/or conclusions,* which complete the graph. (In fact, of the four syllogisms included in our graph, not a single one has all of its elements apparent in the surface form of the Oration, as recorded by

Thucydides: the syllogisms leading to Conclusions 1 and 4 have two of the three, the syllogism of Conclusion 2 connects to earlier parts of the speech, and the syllogism of Conclusion 3 is *totally invisible* in the surface form. Thus – and this is the crux of our argument –, though most of the surface form of the Oration relies on enthymemes, it is only the full structure presented in the above diagram that actually proves Pericles' points in a fully logical fashion. Resorting to the deeper structure is not necessary for the speaker in this case: such a need would arise only if some of the views stated here were contested in argument – something that is not very likely to happen on the somber occasion of a Funeral Oration mourning the city's heroic dead. However, in a differently motivated kind of political speech, with which disagreement might certainly arise – and Athenian political life was full of such occasions –, the orator outlining a surface form such as the one in Pericles' speech (a linearization) would have to be ready, in order to support his argument, also to bring the hidden structure to the surface, at least partly. Some rhetors might not do that until their views were contested. But others, addressing topics that were bubbling with controversy – as famously happened, for example, a century later, when the issue of joining or not with Philip of Macedon was fiercely debated by oratorical giants like Demosthenes and Isocrates –, might take for granted the existence of contrarian attitudes, and thus, playing their own devil's advocates, mostly in the form of rhetorical questions, develop a style that exposed to a large degree – before they were asked to do it by their critics and detractors – the non-linear infrastructure of the arguments.

And this is also what mathematicians do, or rather *have to be able to be ready to do*, when they are challenged by a colleague who is less than convinced by one of their proofs, like the gentleman on the left in the cartoon on page 54: to substantiate their arguments, they have to refer back to the graph in P-space, revealing the elements of the non-linear structure of the proof. And of course, the more mathematicians have to do this, at first for university teachers or advisors during their years of study, and then also for their peers, the more they learn to be their own severe critics, and to internalize a constant – though often unexpressed – dialogic interrogation of their definitions, premises and logical processes, which leads to high standards of rigor. (Should they ever be forgetful of these higher standards, the mathematical community, either in the form of informal dialogue or via the avatars incarnated in the peer review process, will call them to order.)

This process of justification, whether it is internal or externalized, always refers back to the graph of the proof, bringing to the surface as much detail or non-linear structure as necessary. And it is a variation of this process that could have led, quite naturally, to the birth of the deductive method in classical Athens. The example of Pericles' Funeral Oration shows that a politically – if in no other sense – cultured Athenian of the mid-5th century, would be very versed in the demands of an advanced demonstrative style, being able to handle the complex transductions between linear and non-linear forms of an argument – for only those to whom such transductions had become a habit could easily understand the logical infrastructure of a complex linear argument, like the one implied in the segment of the Funeral Oration. It was people who commanded such skills, who first turned to discussing the problems posed by the understanding of space, via diagrams. And it was precisely such skills that – when applied to spatial problems and concepts – gave rise to deductive proof.

That the definitions were clear, and the basic premises "evidently true" (for this was the Aristotelian and thus also Euclidean view of the axioms), as well as the fact that a very limited and basically non-ambiguous vocabulary was used, were certainly necessary conditions. But of themselves, they were not sufficient – using the photos-and-movies language we could say they were type-2! What was necessary to complete the picture was an advanced deductive capacity, and the ability to go back and forth between linear and non-linear, such as had been developed and fine-honed by a century of democracy and the rule of law.

5.2 The *very* old story

In 5.1 we made an argument that proofs *descend* from stories, mostly in a process of bringing forth their non-linear structure. Thus, in a sense, proofs are the progeny of stories. But who bore stories? In other words, if proof got its highly sophisticated methodology from a fine-tuning of the narrative intelligence, applied to controversy, how did the narrative intelligence develop, in a way which explains these complex, abstract skills?

Come to think of it, why should we speak of an underlying non-linear structure of stories? Why speak of storytellers and story-listeners existing simultaneously in the linear *and* non-linear worlds, when at the surface all is linear, word after word after word? Why is this double existence cognitively feasible and, even more so, necessary? To those who are not convinced by the Hergé vignette and the dilemma of Snowy on page 21, as well as the discussion of outward-going non-linearity following it – but not necessarily to them alone –, I want to propose a basic model of this double existence and interaction that is obviously much older than storytelling, and absolutely essential for human existence: space. For, of course, space is non-linear (it extends in two and often three significant dimensions), yet we advance in it linearly, step by trivial step[69].

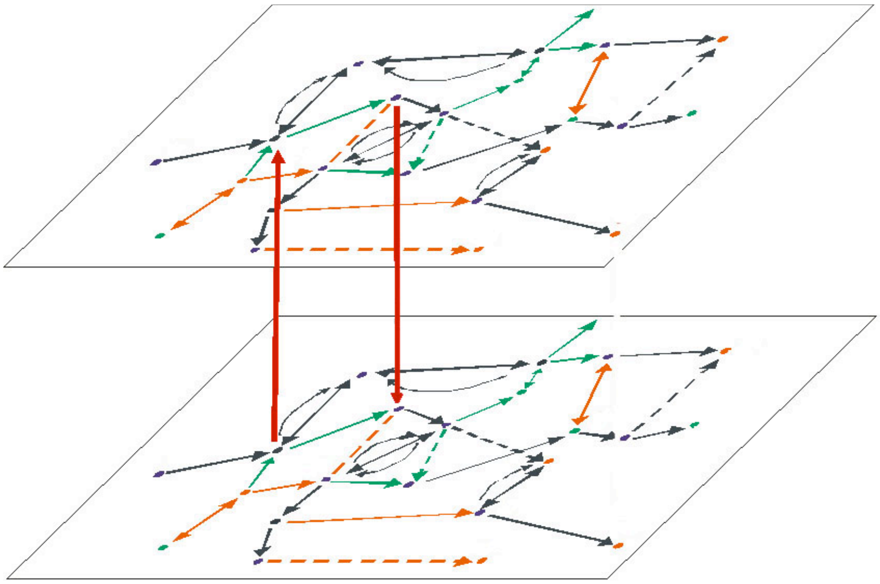
---

[69] For an intriguing old tale of the spatial origins of the *reductio ad absurdum*, see my "Euclid's Poetics", and the discussion there of the story in the *Homilies on the Hexaemeron*, of Saint Basil of Cappadocia, in which indirect proof is attributed to human mimicry of the exploratory habits of the hound.

Our mode of existence in space, and our finely developed skills of spatial exploration, could also underlie the genesis of narrativity. In fact, it is my hypothesis that it is basically the mechanisms underlying spatial understanding, representation, orientation and thus also quests, i.e. the actual, literal-sense, original human quests in space involved in the processes of hunting, gathering, nomadic life and migration, that give the metaphorical ones, in stories, their structure.

Of course, originally space-modeled structures in narrative, as naturally described by the photos-and-movies language, can become much more complex in the process of human cultural evolution, modified and enhanced by the advanced mechanisms of language, chief among which is metaphor, and thus the ability to jump from one quest to the other and continue in a unified narrative[70]. In this case, a story path could jump from one subgraph, in the first setting (e.g. "Hector attacked Patroclus") to a different one ("like a hungry lion attacking a deer, to devour its flesh"), starting in the one, continuing in the other, to effortlessly return to the first ("and pierced his groin with his sword"). The jumps in this case would be the vertical red arrows, the second being the inverse of the first:

[70] These jumps are an argument which might suggest employing the mathematical language of *categories*, rather than of *graphs*, for the photos-and-movies language. Thus, photos would be the *objects* of category theory (rather than the *vertices* of graph theory) and the movies the *morphisms* (rather than the *edges*). The main reason for this change would be that in category theory there is an extensive study of the transitions between categories called *functors*, which historically are developments of the concepts of homomorphisms, i.e. functions that preserve structure. Thus, one can start a trail of movies in category A, which could be the subcategory of a bigger one, and jump at a certain photo to a photo in category B via a functor, in a transition which respects the similarity of the structure in both categories. A lot of modern mathematics uses the technique of creating graphs of such sets of mappings in two, or more categories, and checking whether these *commute*, i.e. whether different paths with the same beginning and ending have similar actions.



These jumps, allowing continuous development from one context (subgraph) to another, can occur both at photo and movie level[71].

But staying at the literal level of one particular physical/geographical space, we can see that the way that this is perceived and dealt with by its inhabitants is very close to a lot we have been describing about proofs and stories. In fact, we can describe this process very well with the photos-and-movies language. In this case:

- *Photos* are locations in space.
- *Movies* are journeys from one to the other.

* * * * *

A person entering a space for the first time with no knowledge of it, will still know the general rules of spatial exploration, having previous knowledge of various kinds of spatial photos (e.g. dark holes can be homes of animals) as well as rules about spatial movies (e.g. they don't go over cliffs) and, more especially, about gradually constructing an inner, non-linear map, through successive linear journeys. In fact, it is this last skill of gradually coming to know the linear through successive non-linear journeys, that is so very pertinent to narrativity – and, of course, proof.

---

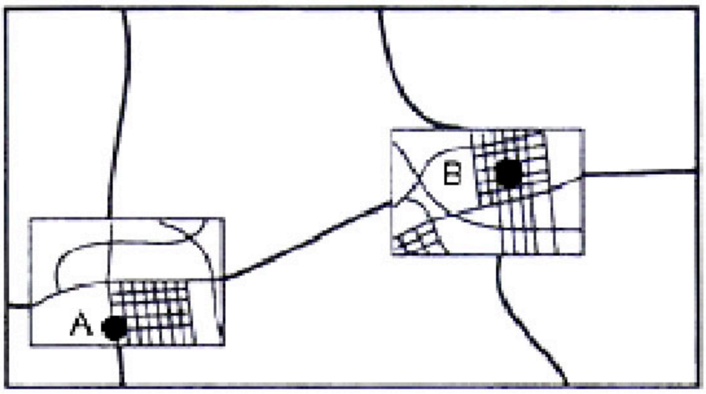
[71] It is my belief, also mentioned in my Mykonos talk as the root of more complex, metaphorical jumps from category to category, that this is a process learned/created in the verbal reconstruction of dreams, many instances of which we meet in their purest form in mythology: the girl turning into a tree is a movie level jump, whereas the half-man/half-bull that is the Minotaur is a photo level. In the Mykonos talk I referred to the earlier instances of this process of joining together different elements, as the origin of the concept of *conceptual blending*, developed by Giles Fauconnier and Mark Turner and discussed in their book *The Way We Think*, and other works. (My favorite is Turner's paper, "The Ghost of Anyone's Father.")

The most obvious kind of movie here is of progress in space, walking or running – and this is clearly analogous to the *t*-movies in stories, or type-1 movies. One might be tempted to think that these are the only kind that exist in space – but not so. Working from the formal model, and its discussion in Section 4, which allows us, in figuring out their type, to disregard any kind of “essence” of movies and just look at their junctions, we can move from the I-junctions of movies of mere “walking” and to spatial V-junctions, i.e. *meetings*. And it is these that create the interesting scenarios (tree-, then dag-like subgraphs) of space navigation that are at the root of narrative, whether these meetings be with human-eaten herbivores or human-eating carnivores, with hunting companions, likely mates, fellow explorers, or hostile members of enemy tribes.

To have a meeting we need *a second intention*, and it is possibly this that projects so many anthropomorphic elements in the landscape and natural world more generally, in early societies. Thus, we don’t only meet a “tiger who would like to eat us” or “a female we want to mate with” or a “hunter who maybe does not want to help us catch that deer but may change his mind if coaxed with a gift of a spear” or “a member of the Bad Tribe who wants to kill us”, but “a tree which will tell us a secret”, a “bird which wants to lead us to salvation”, or “a river which is angry because I didn’t give it any food yesterday”.

Principles like those accompanying outlines, e.g. a rougher or finer sense of a landscape and orientation, *distance* (actually, the name of this principle is borrowed from spatial concepts), *independence of levels* and *self-similarity*, apply especially to space, so much so that it seems very likely that they also derive from it, and the complex ways in which represent space internally. We internalize spatial perception in a non-linear form, from linear journeys, but not always in a simple way of a mere two-dimensional map. The important researcher of the psychology of spatial perception and orientation, Barbara Tversky, prefers the term *cognitive collage* to the more prevalent, *cognitive map*. In a 1993 article she gives as the reason for this term the fact that spatial knowledge is stored “in differing formats, from multiple sources, and from differing points of view” with “collections of partial knowledge (forming) thematic overlays of multimedia and come together to create our spatial encyclopedia”.

In this illustration from Stephen C. Hirtle’s “The Cognitive Atlas: Using the GIS Metaphor for Memory”, quoted in Tversky’s article, we see a mechanism of handling space that we know well from real life.



Starting at a rather fine level of analysis from point A, say a friend's home, we search locally to find the route to the highway, which is traversed with knowledge at a coarser level of analysis, an *outline* you might say – we know that taking the highway is the only way to get from one neighborhood to the next – and then, when we reach our location, in area B, we again search (and thus represent) more finely, to find our exact destination, say our auntie's home where we are staying during our visit to this city. Thus, the instructions "find the highway from A (friend's home)", "go to the neighborhood of B (auntie)", "find the house of auntie" are first at a fine, then a coarse, and then again at a fine level of analysis, a process highly reminiscent of the one in the diagram in Section 3.7 (page 48), where outlines and levels were discussed from stories[72].

Let's go back to the era of the Upper Paleolithic, i.e. from ca. 60,000 to ca. 10,000 BCE, the time of the huge explosion in the cultural life of homo sapiens that most probably also includes storytelling. Since it totally evident that a rather refined mechanism of orientation was well-developed by then, two extra factors are necessary for the development of narrative, the first totally essential and the second perhaps so. The first is the existence of some form of symbolic communication, a form of proto-language, which possibly heavily deictic (with some early Wittgensteins operating on the principle that "that which you cannot speak of you can point at"). The second is intimate knowledge of a particular area, of the kind that can be developed only over a rather long period of settlement and is ideally shared to a large degree by many members of a tribe[73].

This shared, intimate knowledge would give the members of a tribe a rich, common fount of spatial knowledge. This clearly would be a much more effective basis for creating commonly-understood mini narratives, perhaps initially associated with spatial movement, than the images available in a

---

[72] Also, following the concept of Tversky's collage, it's quite natural that this information is also stored internally at different levels, i.e. that there exists no finer rendering, in the traveller's memory, of the area surrounding the highway.

[73] Our hypothesis could be supported by evidence for the transition from a nomadic to at least a semi-stable condition of life for a tribe. Though traditionally this transition is placed around 10,000, the domestication of cattle and the beginnings of farming, there is recent evidence of a settled or semi-settled existence for tribes of homo sapiens that antedates this date by many decades of millennia, in some instances, I believe, going back to the earlier part of the Upper Paleolithic.

purely nomadic existence, in which the ever-changing environment would create a more fleeting kind of mental life, more reminiscent of a dream than a story.

Assuming a good basis of common spatial knowledge, and a proto-language, you can start creating narratives just by symbolically following the paths of progress in space. Thus, “I went to x then to y then to z” is a proto-story, consisting purely of *t*-movies. In fact, at this level stories would be practically indistinguishable from descriptions of spatial courses, either already conducted (reports), imaginary (unreliable reports) or to be conducted by others (instructions).

Assuming the principles of self-similarity and independence of levels to be derived from spatial knowledge, these would allow from day one of narrative skills the much greater flexibility of stories moving simultaneously at various levels, corresponding to the ability to cognitive mix levels of spatial knowledge. Thus, one could say, “from the Big Rock to the Deep Lake, nothing happened, but then as I approached the Tree of the Panther, I saw under it a young maiden. I approached and…” Thus, though the path from the Big Rock to the Deep Lake may be in fact extremely long, say five kilometers, it can be passed over in a sentence, because a) everyone more or less knows what it is and, b) the course was uneventful this time. But when we get to the much smaller area of the Tree of the Panther, the *meeting* with the young maiden (a V-junction in space and thus a strong element of type-2 confluence in the story) makes the speaker shift into a new mode, a finer level of analysis, that is also reflected in the scale at which space is perceived, since the area around the Tree of the Panther may not be over, say, fifty meters in diameter.

Intimate knowledge of a basically constant physical environment clearly functions here as an organizer of the verbal utterances (one is not lost listening to descriptions of actions, because s/he knows what the speaker is referring to, and has an internal sense of the *collage*), thus making both communication and memory easier.

A particularly interesting aside to the last point concerns the ancient technique of *ars memoria* (*technē mnēmonikē*), the first record of whose conscious development and use is attributed to Simonides of Khios. This was subsequently expanded into a codified system – not departing much from the first principle --, elaborated by Cicero and other writers of rhetorical theory, and formed a basic part of rhetorical education in later antiquity, medieval and Renaissance rhetoric[74].

[74] See Frances Yates beautiful *Art of Memory* for the history of the genesis of the technique in antiquity and, mostly, its influence in medieval and Renaissance thinking.

The *ars memoria*, used as the most basic tool for orators wanting to organize big areas of data, is also known as the method of *loci*, and it is based on projecting the new and unknown information on the old and known structure of a well known space (Thus, a speaker will replay in his/her mind a walk from home to the Agora, and "place" in memory, through especially constructed rather eccentric images – for better recall[75] – the main points of a speech that would otherwise be a disorganized bullet list. By rerunning this course, later, i.e. mentally re-walking down the Agora, the images come to mind in the appointed order.)

There is ample evidence that what is proposed as a consciously learned and applied technique by Simonides and his followers is at the basis of many mechanisms of memory. In fact, most people could be convinced of the sense of this by self-observation: this is the same underlying mechanism as that by which, when we try to remember something that was said, or happened or even just thought, we go back, in memory, to the place where we were when it was said or occurred or thought.

Also – and for this, again, there is a strong argument in self-observation – the concreteness of spatial knowledge and spatial courses are often ideal ways in which to organize the much more ethereal dimension of time: thus, when we try to remember what we did or whom we met or what we discussed yesterday, we can do this much easier by first breaking down the day according to the places, the *loci*, at which we were present, in serial order.

Space translates into time, and back, beautifully and this translation is a most powerful tool in our cognitive make-up. But space is concrete, and thus a much better vehicle for recording, structuring and organizing time, via the most-powerful temporal language of storytelling. Were it not for the concreteness of space, and physical objects, and the common reference these can provide to groups of people, it is very unlikely that speech could congeal into anything more permanent than continuous free association of images, with temporality and causality perhaps operating locally, but not globally[76]. And it is this global sense of an axis, and a focus, that makes stories out of action sentences; on the other hand, montages of sequential narratives, without constant focus, can be found in dreams, though a lot of structuring occurs afterwards, in recollection[77].

---

[75] Thus, the "need for grain" might be pictured in the rhetor's mind as a statue of a starving mother, kneeling before, and begging, a well-known baker of the polis.

[76] The article "My mind is like a web browser: how people with autism think", by Temple Grandin (it can be found online) is eminently interesting in connection to this discussion.

[77] In my talk at the "Mathematics and Narrative" meeting on Mykonos, I alluded briefly to how this narrativization of a dream collage could be at the root of passing from a strictly spatial sense of photos-and-movies language, to more metaphorical ones, the passages occurring

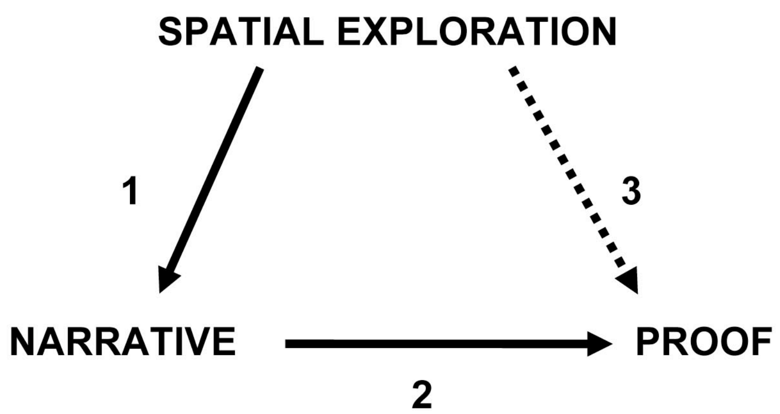
There is one basic difference between time and space however: the first is linear, the other isn't. It is this constant transduction between the two modes, the one of which (time into space) preserves all its information, whereas the other (space into time) doesn't, that is at the root of so many of the peculiarities, and powers, of storytelling – as I hope I showed to some degree, in this paper.

The infrastructure of a photos-and-movies language of space, can also be used – and perhaps this is its most important function – as a generator of new stories, with the mechanisms of spatial graphs operating as organizers, and generators of narrative events: thus, crossroads give rise to decisions, loops mean being lost, a new, unknown road opens the gate to suspense, and mystery[78].

A combinatorics of story ideas, via the use of a proto- and then a more developed languages, could be easily based on the underlying combinatorics of space, which might also account for many grammatical and syntactic structures.

* * * * *

One last point: although we referred in 5.1 to the passage from story to proof, and – in a much more cursory way – in 5.2, to the passage from space to story, there are also traces of this indirect genealogy, from space to proof. In fact, this genealogy need not be only indirect, but could function somewhat as shown in the following diagram:



via processes similar to those Fauconnier and Turner call blending – see also note 12 of this section.

[78] In this connection, it is interesting also to investigate the extensive literature on reading landscape as symbolic form.

Thus, although the original link is 1, once, in the classical times, 2 is also put into place, this is enhanced also by a direct action of 3. There is a rich repertory of spatial and space-related metaphors in mathematician's talk of their proofs ("a way to the proof", "obstacle", "detour", "cul-de-sac", *porism*, "a deep theorem", etc., not to forget of course Turing's trivial *steps*.)

And there is also of course Saint Basil's hound, mentioned in note 10.