Scientific method and game theory as basis of knowledge and language

K. Zawistowski

Introduction

We use method of science (parts of falsificationism) as a foundation of knowledge and language. We draw some parallels to human sensory experience, using recent progress in AI and demonstrate how do we know basic facts about space or ourselves or other people. Then we demonstrate how we can understand and make language with these methods giving examples from Tok Pisin language. Of basic hypotheses of understanding that emerge, one is related to focal point theory, a game theoretical concept. Then we demonstrate viability of this approach for clarification of philosophy. We demonstrate that our theory is good answer to many linguistic conundrums given in late works of Wittgenstein. We improve upon Carnap critique of metaphysics. We comment on Putnam Twin-Earth problem.

It appears that scientific thinking is very important to our lives and also that it can shed much light on apparently unsolvable riddles of philosophy.

As a digression we would like to: - draw some parallels between modern induction and falsificationism, - discuss (as we think) underrated rationality of physics (e.g. OPERA case). - discuss genuine faults of induction (e.g. Replication Crisis) that can be solved by adherence to rigorous inductive logic.

Intro - Sensory data and thoughts

Senses

One could think our sensory experience as something like a movie (a series of pictures, sounds and other sensory data) in our head - we indeed have some sensory apparatus that talks to our brain through electric signals. One could also access sensory experiences of the past, which we often call a part of memory. Accuracy of both of those apparatuses is limited - and memory is way more limited that senses. We percive present and past - sensory 'now' is way more vivid and intense and affecting our senses, compared to memories.

All this sensory apparatus could be adequately replicated with use of electronics

- cameras, microphones, electronic memory, neural nets detecting objects etc - thus it is mostly understood by us as result of some mathematical computation.

Abstraction

We retrieve and process sensory data in a way that is understood and nowadays replicated in various machinery. But we think about this data and make decisions in a way that is not particularly clear. I would call the latter part 'abstraction' or 'hypothesis generation'. This is overlapping but not identical with some concepts in philosophy: e.g. Kant's perception and conception division.

Chollet corpus for abstraction and reasoning.

One paper that helped me a lot is "Abstraction and Reasoning Corpus" by Fracois Chollet, a computer scientist.

He prepared a dataset for testing general artificial intelligence system. It's openly available part consists of 400 challenges, simple games dealing with colored squares on a grid, often no greater than 30x30 squares. Challenge has structure of few examples, showing what question and solution should look like and one test case to solve.

Algorithm to solve one given challenge could be like 'there's red dot and green dot on the grid. draw blue path going from red dot (coming from it vertically), to green dot (coming to it horizontally)'. Or 'there are colored blocks in line with some squares colored, some grey. align squares together so that grey squares stick to each other'.

If a human looks at these riddles, a solution comes to his mind sooner or later. It is not clear how to solve it programmatically in general case. You can figure out, perhaps that most challenges involve squares and lines and points, and it can be represented, but solving most of them requires additional consideration.

Let us denote geometric objects as dot (pixel), line with their coordinates and color in parentheses, colors red, green, blue denoted as r, g, b. Dot $\mathrm{dot}(r, x, y)$ is one red pixel at x, y. Line $\mathrm{line}(b, x, y_1, x, y_2)$ is blue, horizontal line from (x, y_1) to (x, y_2).

Let us consider one riddle. Input and expected output to it is given on Figure 1.

[stuff]

After seeing it and two similar examples one can summarize it as follows: 'There's red dot and green dot on the grid. draw blue path going from red dot (coming from it vertically), to green dot (coming to it horizontally)'. Here's solution - symbolic algorithm to solve this riddle (assuming we give input and output on

Figure 1: Example input and output for one of riddles

left and right hand side respectively).

$$\begin{aligned}(\mathrm{dot}(g, x_1, y_1), \mathrm{dot}(r, x_2, y_2)) \rightarrow (&\mathrm{dot}(g, x_1, y_1), \mathrm{dot}(r, x_2, y_2),\\ &\mathrm{line}(b, x_1 + \mathrm{sgn}(x_2 - x_1), y_1, x_2, y_1),\\ &\mathrm{line}(b, x_2, y_1 - \mathrm{sgn}(y_2 - y_1), x_2, y_2 + \mathrm{sgn}(y_2 - y_1)))\end{aligned}$$

This is example of a challenge and a solution from Chollet's dataset.

The problem is to come up with a computer program that does it on an unknown set of problems similar to 400 problems that were made public.

There was a challenge being held for best solution. The winner scored 20% accuracy on test data set, writing (7k lines long) symbolic solver that worked adequately on known part of problems At the time of writing unofficial state of the art is 30%. At the moment thus it is not possible to replicate human performance on this with a computer program. Given a hypothesis for a solution it is not hard to test it programmatically against examples, but generation of adequate hypotheses in general case is the hard part. This is the reason I think this problem can serve as an approximate model of processing of our sensory input.

Object detection and CAPTCHA

Distinction must be made, between detection of objects and abstraction. Detection of objects in Chollet's corpus case would detect, dots, rectangles, lines, circles etc - rather simple task programmatically. In real world case it can detect cars, rabbits, or trees in visual input or words in audio input. As of 2020 these are largely solved (on par with human performance) problems, with best solutions being neural nets. Neural nets are mathematical equations with millions of parameters. Inputs to them could be e.g. pixel representations of images or sound spectrograms and outputs are answers like 'which word or item is found in the input', given as index or vector with discrete probability distribution. Parameters are optimized on real data.

Let us draw a line (border) between mechanical object detection of state-of-the-art neural network and human object detection. One useful tool of it of it is so called CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apar), a technique often used to prevent programs

(crawlers etc) from access to services that are meant to be used only by humans (such as creating accounts, login). One example of such system is reCAPTCHA by Google. It typically involves solving challenge such as "select pictures with traffic lights". Sometimes 'weird' traffic lights are given - ones that are horizontal or bigger than usual. Also photos are often deliberately of low quality. Traffic light could just look like black box with red dot, nonetheless humans know what it is. There's road, cars and earthwork - and box is just in the right place to be traffic light. This is similar to Chollet-type riddles and also something these neural object detectors are unable to cope with.

Human full object detection is different than neural nets, because it involves abstraction, hypotheses. Human mechanical object detection is similar to neural nets in their function and human visual cortex is possibly similar to neural nets.

Hypothetico-deductive method of falsificationism

Hypothetico deductive method is algorithm most commonly known as method of science. It is often attributed to Popper, influential philosopher of science, but it's origins can be traced to some earlier thinkers as well (e.g. Huyghens)

- We come up with a hypothesis.
- We infer with unlikely or relevant prediction that follows if hypothesis is true and won't happen if hypothesis is false.
- We compare prediction with empirical sensory evidence.

'Relevant' or 'unlikely' here means that prediction should contain knowledge we otherwise don't have - such as prediction of numerical value we don't know (let it be location of time of occurence of something) (For detailed treatment in application to science one could see 'Logic of Scientific Discovery' paragraph 34). We would say that hypothesis is 'corroborated' (Popper's word) if it has confirmed unlikely predictions. That is 'corroborated' is to be confirmed by experience and there are degrees to this confirmation depending on abundance, relevance and variability of unlikely predictions. Abundance and variability are connected, as after one unlikely prediction is done, other similar prediction is not unlikely, because we already expect this this happens. To corroborate is to test by making unlikely predictions sucessfully. If prediction is not verified, a refutation happens, but one could start over with a slightly different hypothesis.

Important distinction to keep the development process scientific is between progressive problem shifts and degenerative problem shifts. If our changes to hypothesis produce lots of new relevant predictions and correctly predicted results - it matters not that wrong predictions and refutations happen, as we are producing new knowledge on average and our final hypothesis is correct and corroborated by new prediction.

If we are proposing explanations for negative results without prediction of new results that's degenerative or unscientific problem shift. This was described in

accessible way in paper by Lakatos [**?**].

Unity of scientific method - induction, statistics and falsification

I would to make somewhat novel point that above Popper's methods and rigorously formulated induction are in practice very overlapping methods and both legitimate to use.

Confusion could be traced to highly influential "Enquiry..." by David Hume and his account of induction, being repeated great many times. Humean induction is following reasoning:

- I see many examples of A implying B. - Thus, always A follows B.

Let it be: I see white swans thus all swans are white. Hume says (rightly) that it can't be justified by reason. And the paradox is that - it is not justified, but we highly rely on such judgements - so let's assume that it is natural instinct - an irrationalist justification.

I answer: no - we don't rely on such judgements. This is artificial umbrella mixing up valid inferences with highly misguided ones - essentially a philosophical failure. Separately - I claim that induction as understood by Solomonoff or Richard von Mises or some solid scientists like Richard Feynman is indeed effective method and very familiar to falsificationism.

For intuive, informal description of what is the problem good source is Feynman's "Cargo Cult Science" essay. It points to "Cargo Cult People" - some indigeneous people of Pacific Islands, who seeing military airplanes bringing supplies, coined up idea that those must be divine messengers with gifts, and proper rituals done by personnel of the airfield summons them.

> "In the South Seas there is a Cargo Cult of people. During the war they saw airplanes land with lots of good materials, and they want the same thing to happen now. So they've arranged to make things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit in, with two wooden pieces on his head like headphones and bars of bamboo sticking out like antennas—he's the controller—and they wait for the airplanes to land. They're doing everything right. The form is perfect. It looks exactly the way it looked before. But it doesn't work. No airplanes land. So I call these things Cargo Cult Science, because they follow all the apparent precepts and forms of scientific investigation, but they're missing something essential, because the planes don't land.

Feynmann point is that modern scientists sometimes are about the same, fooling themselves or other people. Main problem is bringing up so called "evidence" for hypothesis making no effort to look for counterexamples, report possible problems, explore alternatives etc.

For Feynman it appears to be bunch of ad hoc tips made for dealing with practical problems.

> It's a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty—a kind of leaning over backwards. For example, if you're doing an experiment, you should report everything that you think might make it invalid—not only what you think is right about it: other causes that could possibly explain your results; and things you thought of that you've eliminated by some other experiment, and how they worked—to make sure the other fellow can tell they have been eliminated.

We will show that similar principles follow from rigorous formulations of induction, as those by von Mises or Solomonoff. More precisely following 'Humean' induction:

- I see many examples of A implying B. - Thus, always A follows B.

in practice is supplemented by following rules: 1. Complexity of hypothesis must be low compared to data it seeks to describe - simple hypotheses are much preferred. 2. Other than that, all hypotheses are equally probable a priori. 3. Negative evidence is vastly more important than positive evidence. 4. One needs to objectively investigate many different hypotheses.

If we look at Solomonoff induction, 1. and 2. are given explicitly as premises (Kolmogorov complexity and principle of equivalent explanations). As for 4, Solomonoff induction uses theoretical apparatus of Turing machine and sequence all possible input codes to Turing machine to make sure that all hypotheses are tested - but this is not possible to do in practice. This on the other hand should not be considered a limit of the method - we can't generate all hypotheses in general and this affects any system of inference that relies on hypotheses.

Instead we can get close to truth in optimal way in the long run if we would be testing more and more hypotheses - that is essentially going from theoretical induction (Solomonoff's) to practical induction - Bayesian statistics (the basics of the latter can be traced back to Bayes and Laplace). However we need to test all possible hypotheses we can - if we skip some classes, that would likely be a mistake - thus conscious effort to look for alternatives, particularly ones we don't like is badly needed.

Point 3. follows from mathematical form of likelihood. For instance, Bernoulli likelihood of knowing certain hypothesis that something happens with probability p is $p^P(1-p)^N$, where N are negative cases (in disagreement with hypothesis) and P are positive ones (in agreement). If $p = 1$ (something certainly happens) and there are negative cases then hypothesis is refuted. If $p = 0.999$ - that is something almost certainly happens - then one negative example counts as much as 1000 positive ones. Thus ignoring, or not actively looking for negative cases makes tremendous difference.

Let us consider example - if we lived 300 years ago and we observed swans in

Eurasia, seeing thousands of swans - we would come to conclusion that all swans are certainly white. If we then go to Australia and see a dozen of black swans our hypothesis is refuted. We could then argue, that swans are white with 99% probability - but that is not best hypothesis. Better one (according to likelihood and posterior) would be: Swans in Eurasia are white, but swans in Australia are black.

Richard von Mises induction relies on two axioms (it is very close to Reichenbach's induction):

- If relative frequency of sequence of random events converges to a limit and limit is indepentent of place in sequence we start counting, then this limit is called probability. - If we do fair bets on the outcomes of these random events it is impossible for us to make any consistent profit, regardless of strategy we would employ.

First point is axiom of convergence, second is axiom of randomness.

Consider a fair, balanced coin - it lands on heads with 50% probability - and these odds are never changed. If we make following bet: we win one dollar for heads and we lose one dollar for tails - it is impossible to consistently make money of this bet. That is good example of von Mises induction.

It could be that coin is not evenly balanced, e.g. it falls on tails 75% of time. Fair bet would be 1 dollar lost for tails, 3 dollars paid for heads.

Axioms can be rather easily violated. If we have asymmetric coin made of soft metal, that deforms as we toss it - it could happen that odds are different after 1000 tosses than they were in the beginning. In that case axiom of convergence is violated. We can't expect that currently observed odds would be same in the long run. If we see that coin is being tossed with heads or tails upwards it is known that it affects the outcome a bit - in that case we can indeed propose a strategy that outperforms fair bet - which is violation of randomness axiom. Thus, bear in mind that application of these axioms in real world is non-trivial, except in case of physical theories (statistical physics, quantum mechanics) and random games.

Similarly to Solomonoff case if we propose hypothesis that something happens with certainty ($p = 1$), seeing small amount of negative examples in a series changes a lot - axiom of convergence is violated (this follows easily from Cauchy's definition of limit) - and it is not clear what to do about it - thus as point 3 states, negative examples change the outcome a lot. Point 2 and point 4 follows from following observations: - If we apply above induction to two competing hypotheses and see both of them confirmed - there's no obvious reason to prefer one or other. - If we see a strategy to outperform fair bets then it at the same time violates axiom of randomness and is a candidate for new hypothesis.

Point 1 - regarding the complexity, immediately follows if our hypotheses become models forecasting certain probabilities - as typically done in statistics and

machine learning. Often it is dealt with a set of heuristics (bias/variance tradeoff.)

We can see a profound familarity of both methods with hypothetico deductive method - there's indeed a lot in common in these systems:

1. Hypotheses that claim certainty are quickly refuted when negative evidence shows up (this can only happen when we do prediction) and, in Solomonoff/Bayesian case accurate hypothesis can replace inaccurate ones on behalf of small amount of negative evidence. This is implicit, practical hypothetetico-deductive method.

2. How to know that von Mises axioms are satisfied in real world? We can't know precisely if axiom of convergence is satisfied, but we can corroborate it through hypothetico-deductive method. Axiom of randomness is satisfied if we can't improve upon existing hypothesis. Thus we need to look for alternative hypotheses as a rule needed to provide H-D justification for our axioms.

3. Complexity tradeoff is related to progressive and regressive problem shifts in falsficationism - if theory grows more complex without increase in predictive power, that's example unscientific shift.

4. Induction is justified by its predictive power in the long run: "it is guaranteed to eventually approximate the limiting frequency, if such a limit exists. Therefore, the rule of induction is justified as an instrument of positing because it is a method of which we know that if it is possible to make statements about the future we shall find them by means of this method" Reichenbach stated.

A major strength of inductive methods is that it can deal with probability, while hypothetico-deductive method alone can only deal with discrete rules. This comes at cost of big complexity and weaker rational grounds - thus it is generally considered that Popper's theory is more effective and provides stronger justification if it is applicable.

Here we would concern ourselves almost exclusively with hypothetico-deductive method, but one could think of application of Bayesian induction in cases, where one needs to consider uncertainties.

Some examples

Having outlined these methods it is not hard now to spot faults in inductive reasonings - even in cases where they are hard to spot as we shorthly see.

Hume in Chapter X of "Equiry" gives argument against miracles.

It works more or less like that:

Miracles are violations of laws of nature. Laws of nature are well confirmed regularities in nature, established by uniform experience and one law of nature is that human testimonies are often valid but sometimes not.

"No testimony is sufficient to establish something, unless it's falsehood is less probable than falsehood of what it seeks to establish". (formulation based on A. Hajek).

I answer that: if there's a regularity and only one violation of it was reported, then probably it is correct to think that way. But if we use this rule to refute many violations then it is not correct. After few dozens of violations it is unlikely that all of these violations were false. A system that after each consecutive violation presupposes that regularity is as unchanged as before would be essentially an useless system made to ignore negative evidence (see point 3). We can indeed establish existence of rare, uncontrolled phenomena, against extremely well confirmed regularity of lack of such phenomena. As long as they repeat a couple of times (meteorites, rogue waves, supernovae).

Where this fallacy comes from? Hume speaks of comparing probabilities, but as such notions were far too primitive in his times he can't distinguish probability that something happened given law is true from probability that law is true (e.g. whether von Mises axioms are still satisfied).

Another example of various inductive errors assembled together is so-called "Replication Crisis" in science[**?**][**?**], that is growing concern that large part of research findings in some fields (sometimes even most of them) are false. It turns out that large numbers of published scientific studies methodology cannot be reproduced. Paper by Dacrema et al, from field of machine learning, demonstrates how fake, excellent results can be produced by fallacious research that seeks no alternative explanations of these results by double checking all the assumptions needed for producing results (Rule 2 and 4).

Ioannidis demonstrates replication failures in the medical research - claiming that they could be causes by faulty assumptions on probability of a study finding a true result. The fault here is obviously not about irreproducible trials happening sometimes, but happening in such big numbers and producing confusion and loss of time and money that could be avoided. If scientists "leaned over backward", paid attention for negative evidence as a rule and axioms of this probability were tested on regular basis, this would be nowhere near as serious.

Knowing things

Philosophers are in disagreement over issues such as "do objects really exist", "are we in simulation", "are there other minds". We will show that some underlying assumptions of our actions in the world are extremely well corroborated empirical hypotheses, with use of H-D method. (At least, not at most. I do not exclude that stronger guarantee may exist.)

Discovery of our physical existence

Let us assume that we can produce certain variations of sensory input. We can move our hands or our legs, seeing their pictures being moved. This implies being able to predict very well, where they could be for example in next few seconds. My hand will grab a cup if I choose so, or it will wave before my nose if I choose so, or it will remain at rest.

Relevant **experiments** are special case of relevant predictions - they imply relevant predictions that presuppose circumstances we are able to control. It is intuitively 'obvious' to us that our limbs work in certain way, but rationally it can be substantiated as result of predictive corroboration we can and do repeat many times in diffent circumstances (different pictures of our senses).

I did not clearly introduced will or choice - it is commonly agreed we can't 'fail' to will something (Wittgenstein P.I. - 612 onwards), thus willing is always possible to us and if we will to move our hand, moving it is almost always possible (unless one has neurological disease for instance).

Similarly our eye movement does certain movement of our sensory view - we can point eyes in different directions, see our view being transposed over a larger seen area. We can predict part of what we will see a second after our eyes inspected given place.

A reflective surface such as mirror or water or even monitor displaying camera input - allows us to know more about yourself. When we look at yourself in it (not knowing what it is yet) we can quickly learn few things. We stretch our hand - if we see back of it - we see palm in the the mirror, if we see palm, we se hand back in the mirror. Any movements we make are seen in the mirror, any scene behind us is seen in it too, thus we can Apredictively corroborate hypothesis that this thing allows to see anything in front of it. We can see that our eyes and face are about the place where we would expect to see apparatus related to our senses and that our whites move precisely with our sight. To see here means 'first observe, then abstractize, then corroborate'.

We can 'move' (altering our vision in certain way) with use of legs, mostly. Then we can corroborate that we navigate through a space like 3D Euclidean geometry. For example seeing shapes of a cuboid room we can predict what we will see when we move through it. Seeing a cup we can watch it from different angles, then lift it, put it somewhere else and see it stay there - corroborating hypothesis that there are solid objects (as we understand it).

Thus we can go from seeing flat pictures to awareness seeing space.

Kant argued that "we must have knowledge of space independent from experience, we couldn't aquire it from experience". I think not, it can be established by experience extremely well - much better than Faraday law.

Bounded hypotheses

Some of these hypotheses will involve boundaries sooner or later. It is not always true that we will see same thing if we look again in given direction after a time. Some things move or change and description of this will require further hypothetico deductive study. Similar approach applies in science - in XIXc people were resolving deviations from Newtonian model of Solar System. For example Le Verrier used perturbations of Uranus orbit, to predict that there's another planet - Neptune. Anomalous movements of Mercury weren't fixed that way - they were explained by Einstein's General Relativity, a competitive theory. Thus a counter example was made against Newtonian gravity and dynamics. It doesn't mean that Newton gravity was abandoned - on the contrary - it is used much more often than General Relativity.

Is then Newtonian theory true? Not as it was understood in XIX century, perhaps. But we know that these Newtonian equations are extremely accurate as long as we deal with, for example, objects on Earth surface with mass small compared to mass of the Earth. One could say that Newton theory was falsified, but bounded Newton theory reemerged (and was recorroborated).

It is somewhat different in philosophy of science, as focus is made on development of science, regardless of e.g. almost all engineers who keep using Newtonian dynamical equations, not those of Einstein. Popper left some confusion of how 'refutation' works, and Lakatos ("Criticism and the Methodology of Scientific Research Programmes") argued for progressive replacement of hypotheses if new hypothesis reproduces all predictions of old hypothesis and adds new evidence against old hypothesis.

Background assumptions

A handful of problems apparently emerges. Hypotheses that were established in the past are being used at certain point in the future. One hypotheses may depend on other, or corroboration of one hypothesis may depend on validity of many other hypotheses. What if the latter are wrong?

It turns out complex hypotheses depending on other hypotheses can in practice be tested as long as they have predictions and as long as we can test the assumptions too, though their ability of refutation suffers. This is related to Duhem-Quine problem in philosophy of science - but argue that this problem is not a serious obstacle.

Science example 1 - OPERA neutrinos

We will look at some science examples. Reason is that it is hard to find something more convoluted these days than big particle physics experiments with a hundred or a thousand of people working on it and discovery of new physics relying on validity of a large body of old physics. One of such experiments, OPERA produced in 2011 highly anomalous result of faster than light neutrinos. If

controllable beam of superluminous particles was real it would be a theoretical problem. On paper it would allow us to send information backward in time, for instance, thus we would prefer to cap speed of neutrinos at speed of light. Second, it was not confirmed by other measurements of neutrino speed.

What could physicists do in that case? Apply H-D method of course. OPERA people went on testing every bit of electronics and found some fault (failure in time synchronization electronics + loose optical link). Fixing machinery, electronics or software or cracking code is often done with hypothetico-deductive approach as well. When we see issue in program, then we can test isolated parts, predict that they should satisfy tests and narrow down on anomalies. After we fix the issue we expect that whole system will work in use case that was not working, which is another prediction.

If, out of sudden, some elementary physical law needed to make this work stopped working, procedure would be exactly the same. Let us consider a thought experiment: consider that we do mass spectrometry of some bit of matter with accelerator producing ions, magnet deflecting them, electromagnetic measurement apparatus and certain set of predictions of the measurement - and also background assumption such that there's reality being modelled as 3D Euclidean space, there's Lorentz law governing forces acting on particles in magentic field and slow-enough particles move according to Newton dynamics. Out of sudden a tiny, constant, space-time deflection forms up and makes our particles go in different direction - and we get failed prediction, not knowing why it happen. It looks like a problem, but if we keep reiterating our H-D method we can sooner or later find on the issue. For instance if we try doing experiment at different place it would work as expected. So apparatus is broken, we could hypothesise. Then we would keep testing our apparatus - e.g. electromagnet - by moving cathode ray tube in magnetic field. We would find place where deflection is not quite right. Moreover - anomaly would persist when we turn off electromagnet and all particle trajectories would be subject to it, regardless of the charge. Actually photons would be affected too, thus possibly we would be able to see it directly - a weird 'lens' floating air. And if Euclidean geometry suffered a global breakdown it will be seen in other experiments. Sooner or later we would find the issue, regardless what it is, as long as we can keep iterating.

A side note: OPERA neutrino anomaly of 2011 was ocassionally attacked as some kind of huge fault. I think it can be used to demonstrate that researchers (at least in experimental physics) most of the time act rationally, with use of hypothetico-deductive method and getting apparently suboptimal result doesn't imply irrationality. Summing up

1. If after refuted prediction or a surprise, researchers further apply H-D method, to narrow down what is the cause - this is often overlooked in methodological analyses especially done by non-physicists as no one writes this in textbooks or papers.

2. This investigation is often limited by time and budget and also accuracy and

capability of equipment, at no fault of scientists or fault of the method.

Science example 2 - Tycho's Stellar Parallax

Thus it sooner or later happens that we can't immediately investigate further - for instance in case of rare phenomena, that can't be controlled. Or when we don't have any means for further investigation. One of examples of H-D method sending scientists in wrongish direction is Tycho Brahe's attempt to see solar parallax, a result he inferred from Copernicus theory.

What is solar parallax? If we look at distant celestial object from Earth and Earth goes around the Sun, then we should see this object at slightly different direction on the opposing points of Earth's orbit (in June and in December, for instance). It is indeed true that we do, but the difference is small for distant objects (we know that parallax of nearest stars don't exceed 1 arcsecond) and Tycho Brahe couldn't see it. It is instructive to look at the inference Tycho Brahe did with this result. First, seeing that stars are about the same size, he conjectured that stars are about the size of a Sun (we know now it is mostly accurate - Sun is medium-sized main sequence star). He then compared size of a star to size of Sun and deemed that a) either stars are like Sun or somewhat larger and should be close enough to see parallax b) stars are extremely big and extremely far compared to Sun. Considering that the latter was rather problematic and considering obvious sensory experience of stillness of Earth (relationship between atmosphere and gravity wasn't yet understood), he concluded that Earth is still. What's the catch here? We are unable to see the real size of stars. What we see as their size is in fact the size of diffraction picture. If our eyes were better we would see them as really tiny. Of course theory of light wasn't formulated yet either for Tycho to know that - so he picked up simple, well corroborated explanation to his data.

As we mentioned Solar parallax inded exists, but it is so small that Tycho's instruments couldn't find it. It was discovered in XIXc. It is sometimes being look at as an argument against the method, but I would not agree - people got it right in the end, when equipment got better. No one said that this is immediate solution to all the world's problems - like not having good enough technology.

Summary

Based on these examples, I think it is acceptable to have background hypotheses when coming up with new hypotheses or testing them. If result is against what we predicted, we can still just keep testing and narrow down on what failed. Limits to this approach are limits of our capability to investigate - such as equipment accuracy in Tycho's case.

We can pressupose that physical real world exists, or that it is accurately described by Euclid's geometry (within our domain of interest) and use it in new hypotheses - for example that massive bodies fall, that liquid spills on flat surfaces, that night and day come one after other. Empirical evidence for these

pressuposition is on par with anything we could seek to ever establish by H-D method.

All the previous observation don't require any application of language on behalf of mind in question - hypotheses are sets of thoughts in mind. When I call something 'Euclid geometry', is to refer to a concept readers often know, instead of deriving it from scratch. We will investigate emergence of language next.

Other minds and language

Problem of other minds in philosophy can be stated as "do other humans I see have minds similar to one I have". I don't see their minds in same way I see mine. I see their behaviour or I hear their talking - that's it. After our 'mind' figured out how 'he looks like' (as we would put it), with use of mirror surface, he can move on to meeting creatures similar to what he has seen.

After he sees some, he don't knows what they are. But reasonable working hypothesis would be 'they are like me' (we will refer to it as **TLM** hypothesis). They can't talk though, there's no language yet. I think that they can exchange some signs meaning anything like 'positive, pleasant, happy, agree' or 'negative, unhappy, sad, disagree' with certain mimics and behaviour - the former we will call **P-sign**, the lattern **N-sign**. Mind can see, with use of mirror, that these correspond to certain states of thoughts they reinforce (if we smile willingly, we feel somewhat better), and are produced by these states, unwillingly. This can be thought of special case of sensory IO (input-output).

From TLM, and hypothesis that they want to 'communicate like I do' (we will call it **CLI** hypothesis), we can derive inferences sort of "what would I think when I saw them pointing fingers at a rock" and thus predict 'what they will think when I point fingers at rock'.

Source of idea can be found in game theory.

(Schelling point) - in absence of communication, people can often concert their intention or expectations with others if each know that others is trying to do the same"

Quote from T. Schelling who investigated such problems and found people can solve questions like: "three players play game, where they need to put letters A, B, C in order and reward is given when they write same order" (Most people will write 'ABC').

This can be applied to finger pointing and uttering sounds. I point finger at rock and say "rock". It can be derived that given TLM and CLI hold- I must mean something about rock, not about tree that is next to it or about sky above us, as I make rock 'special' by pointing at it. They use fingers to point, too and their senses focus on moving objects more than still and focus on voice more than silence or ambient sounds. They know that I know that they know that.

One peculiar detail on that: act of raising finger could be pointing, or could be doing sth like Socrates

Painting “Death of Socrates” by J.S. David

. He “points“ in one direction, but his face points in other. If we point with our face and eyes and finger - then we are really pointing. Eyes are most important - from mirror we established we use them to look at things - thus our new collegues can think ”he is looking at something“, ”perhaps I should be to“

Problem why finger pointing is done in the direction of finger and not for example other way around comes from Wittgenstein (I think)

. A lucky thing is that our eyes allow it - they have whites - we can thus see where one is looking. There are experiments in psychology that work as follows: if a person on a crowded street stands looking upwards, other will start doing that too

observations taken from J. Peterson lecture online

Thus, mind meets other ’apparent humans’ he can exchange these signs and also he can point fingers on objects and utter sounds, hoping they will understand on behalf of TLM hypothesis. Predictions can be inferred, for example: I will send them P-sign - if TLM holds and CLI holds, they should reply with P-sign. This is similar to a handshake often used in digital communication - ’agreement’ that connected parties will talk in certain code.

Hypothetico deductive language creation on example of Tok Pisin

Wittgenstein “Philosophical Investigation” in first few dozens of sections investigates process that could be used by babies to learn language involving pointing at items, just to conclude that it does not work in obvious way of naming things in the world. Yet, Schelling point would tell that it should work somehow. I think that answer may be provided and with real world examples.

Not with babies this time, I consider babies rather unhappy example. They are notably different in their thinking than adults, no one remembers almost anything from the time he was a baby, they learn language as it is without participating in its development. I will deal with them later.

Better way is to start with languages that were made from scratch not long ago (few hundred of years or less). One such language is Tok Pisin, loosely based on English and used by Papuans and other Oceanic islanders. It came into existence as a result of contacts of these people’s ancestors with English.

We don’t know how this worked but we can look at real examples and try to explain how it could be like.

Tok Pisin is a member of broader category of Pidgin languages, often made between two groups of people trying to communicate without common tongue or means of translation. Russenorsk, an extinct Norwegian-Russian pidgin is another example.

Let us imagine four people who don't share a common tongue. However two pairs of these people can talk to each other. They are well aware of P-signs and N-signs as they talked to people already a lot. They can also point fingers at things and utter sounds, let it be words of their mother tongues. Let us assume without loss of generality pair will be called English, and talk English, and the other will be called Islanders and they don't talk English (without loss of generality) and English would be teaching English words to Islanders (as that is what mostly we ended up with for Tok Pisin, for whatever reason).

Word 'me', 'you', 'he'

One Englishman can point finger at himself and say "Me". What it is, thinks islander: "My name is 'Me"', "I am from people of "me"', "his word for 'chest' is 'me'", "his sort of pale people is called 'me' "

Second and fourth hypothesis can be corroborated as follows: Islander points finger at other Englishman and says 'me'. But he gets N-sign from first Englishman - thus hypothesis refuted - these are not people of 'me'. So it's something in this guy, his name, his body part. Islander then points on first English, about the place on his chest he pointed himself, utters 'me'. N resonse again is received, not this guy, not his body.

Two English now see they can clarify. Second Englishman points at himself and utters 'me'. He then points at first guy and says 'me' with N-sign (not me).

Now it is clear, 'me' is when individual points at youself. Islander can initially corroborate, by calling himself 'me' and receiving P-sign from English.

Similarly word 'you' can be established - one English person points at everyone calling them 'you', but showing that he is not 'you' himself. Another Englishman follows. "He" can be established when facing one person but pointing to other.

Numerals

Numerals can be explained as follows: English shows one finger, says 'one', shows two fingers, says 'two' and so on, up to four. Clearly 'one', 'two', 'three' don't mean 'fingers' in general, it could mean words referring to 1,2,3 fingers specifically, but that would be highly redundant concept. Then, English picks few stones and again, one stone is 'one', two stones is 'two' etc.

Islander can now hypothesise what 'one','two','three', 'four' is, and he can corroborate it by showing English two sticks, uttering 'two' and getting P-sign from English.

Then English can show finger and say 'finger', pick stone and say 'stone' etc. Then, he can explain plurals: 'fingers', 'stone', 'four'.

Living entity

English can point at all people involved uttering 'fellow' for each person. Islander can repeat experiment to see that word doesn't change meaning with regard to the person who uses it. And he can establish with English that rocks or trees aren't 'fellow'. Is then fellow a human, a living thing, a male, an adult? Islander points at tortoise as 'fellow', he could get P-sign, as tortoise could be 'fellow' in vernacular English. Then meaning of 'fellow' would be something like living entity to the Islander. If he gets N-sign it would be human or something more specific.

Word 'We'

Let us think about English world 'we'. Surely it can be explained by pointing and hypothesising as above examples, but it probably won't be. 'We' itself doesn't mean much without a context of conversation 'we - four people here', 'we - two English here', 'we - the sailors on this ship', 'we - the people of England', there's not much reason to remember it.

Now Tok Pisin expression for 'we' is 'mitupela' - it is derived from phrase 'me, two fellow', or 'me, three fellow'. It is made from concepts we discussed and it is way more exact than 'we'. Moreover, an English speaker would easily understand what 'me PAUSE two fellow' means. Me, two people. Like me and other person? Or me and two people, thus three? He can clarify on that responding 'three fellow'. Response would be 'two fellow' - two people including myself.

Analogies and metaphors

Associations in Tok Pisin done with word 'bilong', which comes from English 'belong'. In our example, after teaching few nouns, like shirt, stick, ship Englishman can say 'shirt belong me', 'me and he belong ship' - pointing at ship, shirt, himself and his collegue. And it can be clarified with further examples on both sides, as we did before.

Having this word a lot of expressions can be made with use of analogies and metaphors - and this is what happens. Prince Phillip once remarked he is referred to in this language as 'fellow belongs Ms. Queen'. Knee is 'screw belongs leg', hairs is 'grass belongs leg' (This is not real Tok Pisin, but I write obvious English equivalents for clarity). Continuous tense was made with addition of word 'stop' - which could be another metaphor, result of hypothetico-deductive conversation or both. If our hypothetical Islander wanted to say that certain person is sleeping in his simplified English it could be like 'He sleep - stop'. 'He wake up?' response could be. 'No, he sleep, stop'. Many other phrases of Tok Pisin can be thought of as being result of hypothetico-deductive approach either.

Thus, a language can be made with use of hypothetico-deductive method.

Abstract concepts as parts of hypothesis

Notice example with numerals. An abstract concept is being communicated as part of hypothesis - a quantifier like 'one', 'two', 'three' cannot be explained as a thing. But it is successfully communicated using things. Let us extend it - Englishman shows his hand with three fingers extend and counts downwards - 'three', 'two','one', 'zero'. Zero is fist then? Not really, he repeats with sticks on his hand. 'Three' is to three fingers or three sticks, what zero is to lack of fingers or lack sticks.

I think this hypothetico-deductive approach is really common when teaching children mathematics. What is 'division'? Let us have two heaps of sticks. Each stick from one heap gets equal number of sticks from the other - this number is result. There are three children and nine candies - how many candies each one gets? A question is prediction that student understands.

Why most languages are not like Pidgins - evolution.

Probably most of languagess are not like that, there's much less metaphors or structured phrases, where two or more words together refer to something different. 'Big cat' is one example of a composite - not being big cat on big picture, or big cat as a overgrown house cat - but a tiger or a lion. 'Big cat' is one of few species, that look a bit like really big cat. It is similar to Tok Pisin word for kangaroo (big fella wallaby). Word 'computer' in Finnish is something like 'knowledge machine' - this is somewhat more common for new words.

One very smart feature of Tok Pisin is that humans are good at consciously learning associations - there are memory techniques based on making up associations to whatever information we want to memorize. It could be reason, why it spread and gained a lot of users compared to native languages. Babies are not good at reasoning - they are instead good at intuitive learning of language, so it is possible to teach everyone a language early in their age, regardless whether it is easy language or not.

[baby] Babies learn language intuitively and quickly, adults learn language rationally and slowly.

As for evolution of languages, most of those used in the Europe can be traced to languages used in medieval and ancient era and undergoing slow, gradual changes. These changes include optimizations. Tok Pisin did too - for instance mitupela is pronounced shorter than English source. 'Screw belongs leg' is just 'screw', 'grass belong head' is just 'grass'. This could be done if frequent expression can be shortened without inambiguity - it would usually follow from the context what 'grass' refers too (unless one is a farmer in possesion of a meadow).

Now, if Proposition [baby] is true, one generation of people won't change their language a lot. Some people won't change their way of talking even when those

around talk differently, this can be noticed if one compares older generation to younger. However - babies get a snapshot of language of their parents and have 25-35 years to modify this language and pass it to their children. Thus over many generations language should evolve.

Let us look at two translations of Gospel of John (362 years apart).

" "How do you know me?" Nathanael asked." - NIV.

"Nathanael saith unto him, Whence knowest thou me? "- KJV

New translation easier to clearly pronounce (in my opinion) with 'How' and 'do you'. General composite phrase 'said unto him' was replaced with specific verb 'asked'. Elsewhere words like 'saith' were replaced by 'said' (easier to say) and all 'ye' and 'thou' become just you (distinction is often redundant). One thus could think of it as optimization of communication. We optimize language for other uses, too - as long as we have purpose to pick up words and use them, in place of other words. Some expressions are aestethically pleasing, some are more polite or formal or less polite or formal, some are jokes. Pilots and military have letter pronounciation code: "Alpha", "Bravo","Charlie", "Delta" etc - to pronounce letters clearly. Those could be reasons for the evolution of language.

While Tok Pisin suggests hypothetico-deductive structure, languages that are evolving for much longer time would have lost most of it.

H-D conversations in our languages

Creation of our languages is rather hard to trace. On the other hand I can demonstrate we are much accustomed to hypothetico deductive thinking when using language - which is in agreement with hypothesis that H-D influence was in our language at some point, like it is in Tok Pisin, but was obscured by subsequent changes.

Communication games

People use hypothetico-deductive method when talking in existing language as well. We could want to learn something, to communicate a request, to inquire about information, to double check information. A teller asks client for his ID number - and reads it back to check, if it is correct (as it is easy to confuse). A doctor examines patient - if he suspects lead poisoning - he could ask "Where you work", "do you have contact with lead" etc.

By my experience as a student or teacher I think teaching concepts in mathematics or physics is often hypothetico-deductive, at least if there is small number of students per teacher (a lecture is different to some extent). Often a concept is given in words, with any questions being answered and then few examples of problems are solved, and then student solves problems himself on subsequent classes (with instructor's supervision). What is let's say a Langragian in physics?

After first lecture in classical mechanics you can recite definition, perhaps. After half a year of solving problems you can apply this concept to a new problem.

Adversarial games

There are other kind of language-games - where there's no intention to communicate, but one party wants establish one thing and other party - other thing. TLM still holds, most often, CLI not really. One could call them **adversarial**. Debates, investigations, persuasion, salesmanship and frauds are some notable examples. Some of those involved in these fields use textbooks of so-called "Neurolinguistic programminng" that is influencing other people to do what they want with use of words.

Here's example:

> Tom knew his aunt wanted to catch him in a lie so he said, "Ma'am, I was so hot today that I dumped water from the well on my head. See, my hair is still wet." Aunt Polly was annoyed that she had overlooked Tom's wet hair. "He skipped school again," she thought. She decided to use her secret weapon; before school each morning she sewed Tom's collar shut with thread. "Tom, show me your shirt collar," said Aunt Polly sternly. If he had taken off his shirt to go swimming, she would see the broken thread and know! Aunt Polly was really surprised because the collar was still sewn shut. "Tom, I'm sorry I suspected you of skipping school today. It appears that I was wrong," said Aunt Polly. Tom accepted her apology gracefully. He was just about to leave when Sid said, "Didn't you sew Tom's collar with white thread this morning? Look here." He pointed at Tom's shirt. "Tom's collar is sewn with black thread." "Why, yes, I did sew it with white thread today," said Aunt Polly thoughtfully, and then her face changed. It turned red, her glasses slipped to the tip of her nose, and she screamed, "Boy, that's it!" Tom turned to Sid and said menacingly. "I'll beat you good for that!" he growled. Not wanting to hear what his punishment would be, he ran out the door. "I'll have to punish him tomorrow," said Aunt Polly to herself. "I'll make him work. He hates work, but I've got to teach him a lesson."

(Mark Twain, Adventures of Tom Sawyer).

Aunt hypothesises that Tom skipped school and went swimming. If so - he should have broken the thread. But Tom predicted that already, both her suspicion and her strategy to find out. Thus 'collar was sewn shut'. But Sid predicted that if Tom indeed went swimming he would also prepared some countermeasures against getting caught and looked a bit more closely - and indeed there was a detail refuting Tom's version.

CLI hypothesis doesn't hold here, on the contrary - "They want to obscure truth". Aunt Polly is immediately suspicious when seeing Tom's wet hair - pouring water

from well on his head is hardly worth doing from Tom point of view, she thinks - thus it is an excuse.

One could compare it to mini-max algorithm used in games like chess. Computers play this game with variations of this algorithm, core idea is to hypothesise that opponent will make a move that would be most beneficial to him in the long run, knowing that I would pick moves that are most beneficial to myself. Computer thus would find a move such that e.g. in next 10 moves it would end up in best situation, on the assumption that boht it and opponent choose best moves for themselves (look-ahead is limited by processing capability).

Late Wittgenstein theory of language

Ludwig Wittgenstein in his later publications ("Blue Book", "Philosophical Investigations") came with influential system of philosophy of language.

In this system words have no clear boundary of meaning, but meanings of a word forms a convoluted graph of relations of similarity. Meaning is determined by the way given phrase is used. Philosophers should abandon looking for generalizations, instead just look at uses. Words are like tools he says: a hammer, a chissel - they have certain uses - similarly words are defined by way they are used. He gives example of game: game of chess, game of football, solitaire, hopscotch - are example of games. There's no universal abstraction of a game shared among these instances, but some mutual similarity.

One objection to this system (I know it from preface to Polish edition by B. Wolniewicz, the translator) is as follows - we are told to just "look" at word's use, but there's no way how to look, as no theory is given and no theory could be given: (Preface to PI, paragraph 7)

"We are to establish that two words mean same thing on behalf of identity of their use. Very well, but how to recognize it? This critical problem is left totally undetermined in "Investigations"."

This would require a theory (Wolniewicz points out) - but Wittgenstein doesn't give any theory, just bunch of loose comments with many backdoors to "sneak out".

I think I can give a theory that answers to large part of criticism in PI: that language is being developed and understood with use of hypothesis testing, as a mean of communication. As we described: CLI, TLM and FPT are most important hypotheses being used. - Words are related to sets of pictures ("rabbit", "brick") or sets of ideas ("four", "one meter", "walk") in mind. - Meanings of words are only in close-enough agreement among people, so that communication is possible and efficient. - Universal definitions across diverse group of speakers are impossible, similarly very precise definition of words are impossible. - But at the same time there's a lot of structure in between. A continuum of decreasing clarity, a continuum of increasing generality - as we go from small groups of

people to large groups of people. - Language evolves - if few people establish new term and uses by hypothesis-testing conversation and it spreads virally - it can be accomodated into language - unused terms may be forgotten.

"Pain" examples

Wittgenstein uses examples of pain, to demonstrate issues regarding mostly comparison of our mental states to knowledge about mental states of other people and related deceptive uses of language. Let us examine few examples:

""I grant you that you can't *know* when A has pain, you can only conjecture it", you don't see the difficulty which lies in the different uses of the words "conjecturing" and "knowing"

Let us notice that this is easy to clairfy in our theory. We introduced TLM hypothesis as one of fundaments of our understanding of language. This is one example where this is apparent. Of course we have direct sensory experience of pain (for instance) in our body and knowledge of reactions it produces in us. We don't have direct experience in other people, but very well corroborated TLM hypothesis tells us that it should work the same in other people. If knowing that "I am in pain" is knowing state of sensory input it is somewhat better way of "knowing" than the latter. We have direct access to sensory input in the first case, we do inference from background hypothesis and sensory input in the second case. There's indeed a difference in level of certainity, but I think it is vast exaggeration to make a distinction between knowing and conjecture - TLM assumption is very basic and essential thing. We commonly use phrases like "I see him smiling", "I know he's angry", "This person" - they all presuppose TLM.

There are other considerations: cases when someone may pretend being in pain: a pupil who doesn't want to go to school and wants to convince his mother that he shouldn't - for instance. In such a case we would differentiate that accordingly in our language: "I think he is really in pain" - "I think I know he is in pain".

"If I point to the painful spot on my arm, in what sense can I be said to have known where the pain was before I pointed to the place? Before I pointed I could have said "The pain is in my left arm" "

As we argued we know (as well confirmed hypothesis) that we are able to move our body and we receive sensory input from it - temperature, touch etc. We can pinch our hand or leg, produce feeling of pain - then corroborate by predictions that it produces predictable sensory input. Then if we fell pain in our tibia, we can relate it to closest feeling of pinch we can produce - this allows us to point our shin. Of course our brain does it somehow more efficiently, I am deriving a plan to know it rationally.

"It is conceivable that I feel pain in a tooth in another man's mouth; and the man who says that he cannot feel the other's toothache is not denying *this*." " Suppose I feel a pain which on the evidence of the pain alone, e.g., with closed eyes, I should call a pain in my left hand. Someone asks me to touch the

painful spot with my right hand. I do so and looking round perceive that I am touching my neighbour's hand (meaning the hand connected to my neighbour's torso). " "Ask yourself; How do we know where to point to when we are asked to point to the painful spot?" 'I said that the man who contended that it was impossible to feel the other person's pain did nor thereby wish to deny that one person could feel pain in another person's body- In fact, he would have said: "I may have toothache in another man's tooth, but not *his* toothache' 'Thus the propositions "A has a gold tooth" and "A has toothache" are not used analogously. They differ in their grammar where at first sight they might not seem to differ.'

Wittgenstein in this paragraph argues that seeing that person A has gold tooth is much different than seeing A has a toothache, and yet different from having a toothache in A mouth - and rightly so. Let me explain it - in first case we see, through our sense of sight - a person (TLM used here) with a golden bit in his or her mouth. Second one (A has toothache) is just special case of other person's pain we dealth with already. Third case (I have toothache in A's mouth) is yet different. I think we can't comment on intuitive perception of such thing - but we can indeed know rationally that such pain exists. All it takes is to take a steel dental probe and hit A's tooth (possibly after a lot trial and error to find where pain comes from) - if it immediately intensifies our pain (per analogy to teeth in our mouth) we would know that we have pain in A's mouth. Other pain examples (feeling pain in furniture and such) can be clarified in equivalent way.

Ostensive definitions

First few dozens of paragraphs of "Philosophical Investigations" deal with ostensive definitions - the general point is that they don't work - or not necessarily work.

(PI 2) introduces situation (language-game as he calls it - a situation of practical use of language) when two construction workers communicate as follows: Worker A yells "block¡', "slab¡', "pillar" or "beam", then worker B brings him specified item. Then extensions are given in subsequent chapters - worker A may yell "a-slab" or "d-slab" to obtain some specific type slab or "this-there" where he points to some item and some place to carry this item to. Then in (PI 10) he makes point that meaning of words must be defined by their use. This is probably argument against theories, that would suppose that "slab" is attached to a picture of a "slab" in mind (such as one advertised by "Early Wittgenstein" in Tractaus).

I think he is mistaken to derive support for it from this example, and it can be demonstrated by his own method of looking at such situations in ordinary language. I think they are not uncommon - construction workers may have good reasons to communicate by yelling a handful of words - for instance if someone talks to you from few floors above and next to a guy cutting steel with a grinder, you are hardly able to hear anything. Thus a simplified language is made of

yelling. In case elaborate conversation is needed, you need to come closer.

How those come into being? One is explicit agreement: they may say: "Look, I will be doing this, and you will be carrying needed items - if I yell 'block' or 'slab' you bring me what I say" - . Regulations on a shooting range typically say something like: "In case of danger anyone should issue command 'STOP!'. If such command is issued no one is allowed to shoot". A soldier, when issued a command let it be "alert!", knows exactly what he should do, according to rules he knows by heart.

One could think of it as context of conversation. "You will be carrying items, if I yell "brick" you bring me brick" and then "Brick¡', "Brick¡'. I think this is compatible with picture-theory - word "brick" referring to a picture of a brick in mind. It is obvious that one should carry bricks if it was specified so before.

One may argue on the contrary it as follows:

1. We can't assume that it is always (in these situation) agreed upon that one would use such and such "commands".
2. Even if we can, same thing may be called "context of conversation" or "definition of word'.

Let me now argue that this explicit agreement is superfluous and that usual definition of word is applicable. I am going to use focal point theory (FPT). A bricklayer may yell to his assistant "Bricks¡', without prior agreement, having intention to get some bricks from him and be almost certain that he will be understood. Bricklayer's job is to assemble a wall or other structure from bricks and assistant's job is to help him, by carrying needed items. Both know that and know that other party knows that as well. So sentence "Please bring me as many bricks as you can efficiently and safely carry" is reduced to "bricks" - other parts are superfluous. "Bricks¡' can be understood as "Please bring me as many bricks as you can carry", not as "Bricks have fallen on my leg", "What a wonderful pile of bricks" or " I don't like bricks a lot" or "Bring me only two bricks this time". Bricklayer, knows what assistant should expect to do does not say what he should do. Assistant knowing that bricklayer knows that, does what he should do by default: "Carry bricks" and as many as he can (to do his job efficiently and safely). Same for "this-there" example - on behalf of FPT and above stated facts it means "Please bring this item (one I point at) to this position there (I point at)" - not "This thing was there before". Same for examples of (PI 21) and (PI 27)

Similarly, if I come to shop, and say "a lighter" it will be understood as "please sell me lighter". If shopkeeper says "one dollar" I will understand it as: "lighter cost one dollar, please pay me one dollar". A ticket inspector would say "ticket" and mean "Please show me your ticket". These are all obvious application of focal point theory. But "lighter" here remains lighter, "dollar" remains dollar, "ticket" remains ticket - meaning of word is not affected - and picture definitions are adequate.

This is direct answer to (PI 19): Wittgenstein says that "Slab¡' is indeed "Bring me a slab¡' shortened, but why not other way around? "Bring me a slab¡' being extension of "Slab¡'. "Slab" could be related to picture of object in mind, "Bring" refers to action, "me" to person. "Slab¡' is "bring me a slab" shortened with use of FPT.

Colour, numerals, length

(PI 29) raises objection that there's no ostensive definition for colours, numerals, length etc. As we already demonstrated numerals can be defined in our Tok Pisin example through hypothetico-deductive discussion. Colors are not different - we can demonstrate concept of blue to other person by demonstrating a couple of blue items - blue glass, sky, sea, ink and also few non-blue items. After other person tries to give us few more examples and receives feedback from us - showing us items in different shades of blue and getting feedback - we can establish close-enough agreement.

It is important to comment here, why 'close-enough agreement' works well in real life. For example if all our examples were glossy blue, our collegue could be not certain whether 'blue' includes only glossy items, or whether it also applies to something mat or rugged. But he would call mat blue as well if only there's no better term for mat blue (or he could call it "mat blue" or clarify when he needs it).

Notion of length can be taught by a more complex set of analogies. First we establish numerals. Then we grab a stick - we call it "one meter" ostensively. We use it to mark off one meter on ground in various horizontal directions and also on a wall - and call all those examples "one meter" too. Then we proceed to mark off 2,3,5, 10 meters. Then we repeat this drill with different template (let it be 20 cm-long stick or 25 cm long foot or 1 meter long stride) - that's how hypothetico deductive process is started. Our "student" can corroborate his understanding by performing measurements under our supervision. Then we proceed to say that "town is 2000 meters away" and so on.

A mistake in demanding what cannot be given

I must stress that those are language expressions correspond to close-enough agreement of ideas or pictures in people's head and this agreement is not exact. Exact agreement can't be achieved, as our knowledge of other people's internal mental states is very limited.

Expecting otherwise entangles us in myriad of unsolvable problems and lengthy discussion of reading in (PI 157-165) is good example of it. Wittgenstein starts with pointing out example of a pupil who "reads" a word and other people wonder whether he really did or just made a lucky guess. If he reads more words then we can see he indeed reads.

L.W. here points at "contradiction": on first word teacher says that pupil doesn't

read, on reading a couple more of words teacher is convinced that pupil indeed reads. But how about this first word L.W. asks: was it read? This question is nonsensical, he answers himself, unless we assume that we start reading.

I am rather surprised how neatly this apparent problem is solved by method of hypothesis testing. Teacher wants to find out whether pupil reads. “No, you don’t read” can’t be conclusive statement of fact - teacher has no access to student internal mental states. Teacher hypothesises “you don’t read”. Pupil reads a bit more and teacher is conviced otherwise. One could estimate probability of guessing content of text by chance as exponentially convergent to zero with it’s length - which is exact scientist way of saying “this is impossible for all practical reasons”. Teacher, in a contrived example could say something like ’you don’t read’ indicating a conviction about such internal mental states - but this can be explained by FPT. Teacher and student both know that teacher can’t state facts inaccessible to him and know that other party knows that. ’You don’t read’ is ’I think you don’t read’.

In general a comprehensive method of reading is taught by hypothesis testing - a teacher demonstrates how to read letters, syllables and words - and then expects pupil to memorize, imitate him and then generalize to new words. If pupil is mistaken, he is being corrected by the teacher. Later they work on other topics, such as understanding of text being read. A hypothesis could be: if pupil understands texts he should be able to pick matching paraphrase. Such exercises are common in textbooks for teaching foreing language. And more importantly - they are important part of exams. How we determine proficiency in language? Student solves bunch of exercises, writes an essay, participates in conversation - that is content of exams.

It is hypothesis testing all over the place in application to what can be externally observed - that’s how people find out what “reading” is in our method. And this definition is shared closely-enough to allow communication. How about internal mental states in regard to reading? No way to know and no reason to assume that they would be same among different people - that’s universal answer to the rest of Wittgenstein’s discourse.

What L.W. means by “language games” is not clear, as it is not clear what distiction he seeks to establish (I borrowed his own argument).

We can think of global definition of words: i.e. whether there’s one meaning of a word for most people using it. If we limit ourselves to one language - pretty obviously it can’t be so - there’s regional variation in the language, there are lingos of various societal or professional groups etc. Thus meanings of words “in general“ will end up being overlapping, convoluted graphs but that’s hardly a surprise and no contrived examples are needed to see that.

Somewhere below language games starts indeed - a group of people engaging in use of language form clear meanings of words. But is there a clear distinction of meaning ’in language game’, and ’across language games’? Like: first is given use case of word, second is convoluted graph without clear boundaries? I think

not - I claim that meanings inside of game are only in close-enough agreement for communication to happen. Outside of given game they could be slightly less defined, but this is a continuum, not a discrete boundary.

Think of "Brick", "Slab" example - I demonstrated that picture theory is completely adequate here - "brick" could be still understood as picture of a "brick" in mind in the language game that seeks to demonstrate its inadequacy.

I think problem here is in desiring two things 1) a grand universal definitions simultaneously valid for at least most of English speakers 2) definitions precisely connected to mental states shared by most people. It is not hard to see absurdity of both by producing few counterexamples, but that doesn't mean that there's no objective structure somewhere in between (as L.W. seeks to establish).

On the contrary, if we assume that word meaning emerges as hypothesis testing phenomenon and it's meaning is agreement of ideas close-enough to allow communication we can explain his counterexamples.

Did N exists

Point in (PI 79) is more or less that: a bunch of people argue on historical person - let it be Socrates - known from written sources. Some believe the narrative that Socrates lived and performed certain things. Some deny parts of narrative. Some of them even claim that Socrates never existed, Socrates wasn't called that way when he lived or biography of Socrates in fact refers to few different men.

Then argument is made that e.g. "Socrates didn't exist" means different thing, depending on what meaning of Socrates is. "Socrates didn't exist", because his real name was Cleanthes, "Socrates didn't exist" because he's a myth.

One way to deal with it [**?**] is to assume that "Socrates" is a category that includes all options. Socrates didn't exist - all kinds of Socrates didn't exist.

Surprisingly - application FPT/CLI produces same result. If I say "Socrates didn't exist" I would mean 'all kinds of Socrates didn't exist' by default (in absence of context) - I reference certain feature to make distinction. My audience reasonably expects that if there were two Socrateses, I would tell them explicitly, on behalf of CLI hypothesis with FPT. There's no way to infer it from "Socrates didn't exist" without context, so it would be inefficient to say this that way - as they could know.

Chair example, pictures, computer games, simulations

L. W. in (PI 80) comes up with following situation: "There is chair" we think. But as we move closer it dissappears. Then it reappears in few seconds. Then he asks whether reader had rules prepared to tell if this is still a chair. I think do have such rules indeed. What is violated is his example is the assumptions about perception of space and navigation through space. I would expect thus that chair is still chair. We would refer to it as chair.

This works that way in various simulated realities we use in XXI century - computer games or software used for design. A chair seen in such software is still called chair. Even less than that - a chair on a picture is called chair. L.W. should know the latter very well, part of (PI II) is dedicated to seeing things on pictures - a rabbit on a picture or a duck on a same picture. A chair in a space that sometimes dissapears is much closer to real object than picture, thus it is called "chair".

Why do we call picture-rabbit a rabbit? First, "rabbit" is best phrase in our vocabulary - it looks mostly like rabbit. Why don't we call it "picture-rabbit" - reason is again FPT and TLM. If we talk to other person in a room, with picture-rabbit on the wall we know to what kind of "rabbit" we refer to and other person knows that as well.

If we were talking by phone while playing computer game and we told that bandits are here - then we could be misunderstood (other person will think about), so we don't say it that way. On the other hand if we played online, talking to people playing with us - we would be understood. How to say something? Efficiently and in a way that other person understands me. How to do that? TLM, CLI, and FPT are here to help.

Bear in mind that our hypothetico-deductive concepts of seeing objects and locomotion through space can be disconnected from any additional notions (what they "really" are) - thus if we play a game with 3D graphics and virtual reality set, we can apply same spatial or geometric concepts. We can even sometimes forget that it is not "real" (it is an excercise left to reader to think what is "real").

Failure to understand

(PI 185) describes pupil who is told to write down elements of arithmetic sequence $a_k = 1000 + 2 * k$ (1000,1002,1004...) but gets it wrong (1000, 1004, 1008) - and can't be convinced that it is wrong.

Indeed it happens that some people won't get more advanced, academic mathematics - since capability and desire to do so varies through population.

But in that case it is at least clear how teacher should proceed with our method. He can do somethink like - "Look, you first wrote 1000, which is good, but if we subtract it from next element we get 4. This is in contradiction to definition I wrote, as $a_{k+1} - a_k = 2$. how to fix that?".

Pupil can insist it is in agreement, or derivation is wrong but in that case process can be continued to show that he denies something basic about math: definition of addition in natural numbers for instance. Why could be that? Perhaps a deficiency similar to colour-blindness - but nothing to do with our method.

Application to problems

We will look at two language problem that can be easily solved with our theory.

'Nothing itself nothings'

"Nothing itself nothings“ is somewhat representative quote from Heidegger. Vienna Circle people weren't fond of it as a philosophy. One could look at, for example "Elimination of metaphysical thoughts by logical analysis of language“, Section 5 by R. Carnap. Few sentences he quoted from Heidegger could be, he claims, fallacious use of word 'nothing' as noun, while it is a quantifier often used in negations ("Nothing can fix this issue" i.e "This issue can't be fixed (by any-thing) "). Then he claims it could also be that Heidegger's 'nothing“ is something else than ordinary use of this word - an "emotional constitution“ for instance. But this is not the case, he claims, as Heidegger starts with "What is to be investigated is being only and-nothing else; being alone and further-nothing“ clearly pointing to it's ordinary use "nothing else“, "and further nothing“. Then quotes from Heidegger are given, on his dismissal of logic and science. Thus, Carnap concludes, Heidegger is, probably, talking meaningless sequences of words.

Can Carnap establish that? One accusation I found is problematic translation of Heidegger who was German and his works are hard to translate: 'being' could be 'what-is', 'Nothing nothings itself' - 'Nothing nihillates of itself' - depending on translation. Nothing nihilates of itself. "Nothing“ is a thing that rids of this thing itself? This still looks to be nonsense as 'nothing' still must be noun. On the other hand thing " that rids of itself“ involves some action and relation like meronymy or ownership being at the same time reflexive (between thing and itself).

First, I am not certain there's no meaning here. Second, the riddles are not easily solved, the riddles multiply themselves when we attempt solution. Carnap points that Heidegger scornfully dismisses logic and rationality - in that case we indeed lack tools to proceed with solution. We don't have any clarification from Heidegger, he never responded Carnap as well.

We can't assume logic to work - that's first problem - we silently presupposed logic everywhere so far. When we read above excerpts Heidegger it is apparent that it could be adversarial language game - like one with Tom Sawyer and his aunt - in that case we probably can't rely on CLI hypothesis for deciphering these. Meaning is obscured from us for some reason and no one knows it (there are competing interpretations).

Based on other Heidegger's works we can think that TLM is violated either. He [**?**] appeals to some sort of mystical revelation, his special inner insight, H. Phillise says:

> Heidegger relies on a epistemic model derived from theology, and

> assumes that he is the recipient of some kind of revelation... (...) What Heidegger counts on, then, is that we will simply believe what he says. He uses a number of authoritarian rhetorical stratagems in order to obtain this perlocutionary effect, and he is remarkably successful in securing it. (...) Because Heidegger in section 7 of Sein und Zeit calls empirical phenomena "vulgar" phenomena, we might label empirical history "vulgar" history. To vulgar history, Heidegger opposes real or authentic history (eigentliche Geschichte), which is the sequence of fundamental stances underlying vulgar history. Real history is "necessarily hidden to the normal eye." It is the history of the "revealedness of being" (...) Heidegger belonged to the elect, to those favored by Being, who were destined to hear Being's voice

Heidegger mental states regarding these matters can't be related to other's mental states - thus TLM probably doesn't hold.

I would rewrite Carnap's criticism as follows: there's no way to know what Heidegger means, as he topples down most basic structures we use to understand language.

Putnam's Twin Earths

Putnam Twin Earth problem is as follows: there are two identical planets with copies of same people having identical sensory experiences of everything. One planet is our Earth. The other is Twin Earth - with only difference being that our water (hydrogene dioxide) is replaced with compound XYZ with exactly same physical properties - both populations refer to their compound as "water". Or when it freezes they refer to it as "ice", when it boils they refer to it as "steam". The date of experiment is set in Earth's medieval era, where neither population knows any chemistry.

Question is "Whether Earthling say 'water' and his twin from Twin Earth say 'water', do they mean same thing". Putnam thinks not, Earthlings refer to H20, Twinearthlings to XYZ.

From our point of view 'water' is just hypothesis for something wet, liquid transparent, sometimes drinkable - thus they refer to same abstraction equally applicable to H20 and XYZ, as 'three' is applicable to 'three sticks' and 'three stones.

I think this is reflected in our language. There is a drink called 'coffee'. It can be prepared in different ways - but one could think that it should be made from roasted coffee beans (Coffea arabica or Coffea robusta). But drink called "coffee" is also made from rye, barley and chicory root - it has similar color and similar taste. It would be called 'cereal coffe' - for distinction, but there's still world 'coffee', referring, probably, to similar experience of drinking.

How about disk? Disk is sport equipment for throwing. Optical device for holding data is compact disc - compact refers to size, probably. It is abreviated

as CD. But same with same size device only slightly better is not 'compact' anymore - it is DVD - digital video disk. Of course you can write both with music or video if you want. Of course CD is digital storage too. So if I have video on CD as it is often understood, I can tell my friend it is DVD. I would not do that, because he would understand it wrong, it makes no sense. If I want to communicate with him I would clarify 'It's video on CD, and you need such and such equipment to play it'.

9 Herman Philipse, Heidegger's Philosophy of Being, cited in David Auerbach, "Authority by Obscurity". "Tragedia werbalnej metafizyki", L. Chwistek, 1932 Ludwig Wittgenstein "Philosophical Investigations" (in Polish), PWN , 2012 Karl Popper, "Logic of Scientific Discovery" (in Polish), Aletheia, 2002 Imre Lakatos, "Criticism and Methodology of Scientific Research Programmes" John Ioannidis, "Why most published research finding are false" "Are We Really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches", Maurizio Ferrari Dacrema, Paolo Cremonesi, Dietmar Jannach