

Woods's Naturalization of Logic: Logic for Real Agents

Douglas Niño

edison.nino@utadeo.edu.co

Draft

Published version in: Dov Gabbay, Lorenzo Magnani, Woosuk Park and Ahti-Veikko Pietarinen (eds.), *Natural Arguments: A Tribute to John Woods*. London: College Publications, 2019, pp. 625–659. ISBN: 978-1-84890-302-9.

1. Introduction: The Project of Naturalizing Logic

The purpose of naturalizing logic is to give a principled account of human reasoning that takes into account the findings of the various sciences of mind. This makes it an analogue of Quine's project of naturalizing epistemology (1969). But, in contrast with the Quinean spirit, it involves a particular pragmatization of logic, which blurs the distinction between logic and psychology:

Logic should investigate the cognitive aspects of reasoning and arguing. If what we have said about reasoning as a facilitator of cognition is so, and if logic retains (or re-engages with) its historic mission as a science of reasoning, then logic must take account of what cognitive agents are like, what they are interested in and what they are capable of. Given that beings like us come rigged with psychologies as standard equipment, the hostility of logic to psychologism is risibly inapposite... logic's toleration of psychologism must embrace the idea of reasoning in the service of cognition and take due notice of reasoners as performers of cognitive tasks (Woods, 2007, p.8-9).

This project does not imply that a naturalized logic should be the only logic. It does imply that real reasoning performed by real agents in real time is worthy of study. In studying such reasoning, traditional normative standards of deductive and inductive consequence may not be adequate. Not even the smartest human is capable of drawing all the logical consequences in a given logical system, for humans do not draw consequences in a logical space but in a psychological one. We are then to distinguish between *drawing consequences* and *having consequences*, for not every consequence a system has is, or may be, or even should be, drawn. Thus, a naturalized logic concerns itself with the phenomenon of drawing consequences. This consideration calls for a further distinction: the consequences an agent actually (or usually) draws are not always the consequences she should draw on a given occasion and we therefore need to discern the normal from the normative. Woods anticipates this and advocates their convergence (the NN Convergence Thesis):

CONVERGENCE OF THE NORMAL AND THE NORMATIVE (1):

As a first pass, and when there aren't particular reasons to the contrary, how we *do reason* from premisses to conclusions is typically how we *should* reason. In other words, in matters of consequence-drawing there is a trending convergence between the *normative* and the *normal*, between *what is usually done* and *what is rightly done*" (Woods, 2013, p.52).

THE CONVERGENCE OF THE NORMAL AND NORMATIVE (2)

When there aren't particular reasons to the contrary, it is typically the case that how we reason from premisses to conclusions in real life circumstance is either *accurate* or *apt* or both.

COROLLARY: In the matter of drawing consequences reasoning rightly is more like breathing rightly, and less like interpreting quantum phenomena rightly.” (Woods, 2013, p.54)

Take it as given in the absence of particular reasons to the contrary that humans reason well when they reason in the ways that humans normally reason in the conditions of real life. (Logic Naturalized, n.d. p.17)

If the mere juxtaposition of the words “logic” and “naturalized” may irritate many formal logicians, the convergence of the words “normal” and “normative” may cause them severe urticaria. But this need not be so if we see human reasoning on an evolutionary basis: had the normative not become normal, the human species would have become extinct.

What we mean by “normative” nonetheless requires further explanation. The standard of normativity in classical logic (as in Kantian ethics) is very high, but what exactly does it consist in? On the other hand, if logical methods have traditionally been associated with the production of knowledge, what does this naturalized logic have to say about it? Again, Woods anticipates the question. He puts forward a particular form of *reliabilism*, the Causal Response Model (CRM), opposed to what he calls the Command-and-Control Model, associated with the traditional idea of knowledge as “justified true belief”, in which acquiring knowledge involves ‘case-making’. In CRM, the “justification” part is removed and replaced by a causal component. For the most part, reasoning occurs “down below” (Woods, 2013, p.122), that is to say, in an uncontrolled, automatic or semi-automatic, unconscious way. Reasoning, of course, also occurs “up above”, that is, in a controlled, deliberate and conscious way, but the latter is much less frequent than the former. A naturalized logic therefore requires a particular epistemology. According to Woods, the CRM is compatible with and appropriate to the NN Convergence Thesis: “for the most part, a human agent’s premiss-conclusion reasoning is the right way to reason when his conclusion-drawing mechanisms are in good working order, and at present working in the right way, engaging good information in the absence of hostile externalities” (Logic Naturalized, p.17). And regarding knowledge, Woods specifically says:

A Causal Response Description of Knowing: A subject knows that *a* provided that *a* is true, he believes that *a*, his belief was produced by belief-forming devices in good working order and functioning herein the way they are meant to, operating on good information¹ and in the absence of environmental distraction or interference (Woods, 2013, p.93; the underlining is mine).

In this paper I would like to contribute to the project of naturalizing logic in the spirit of Woods’s work. In what follows, I devote a section to each of the three expressions underlined in Woods’s quotation above: subject, believes and good information. In the first section, I advocate a re-examination of the agent-agenda proposal advanced by Dov Gabbay and John Woods in 2003 and 2005. This will have, I hope, some bearing on the CRM and the NN Convergence Thesis. In the second section I examine Woods’s characterization of abduction, which is a ‘privileged case’ for third-way reasoning, and in this sense, for real, everyday reasoning. I venture here (with the help of Peirce’s writings)

¹ Information is good for belief just when it is accurate, current, and complete (Woods, 2013, p.114).

to make one additional clarification to one of Woods's theses about abduction: what has to be done with the abductive conclusion, once it is obtained. This reflection on the abductive conclusion is also associated with the second premiss of the abductive schema, which, in turn, poses two of the most challenging problems for a naturalized logic: plausibility and relevance. The third section is dedicated to them. On the one hand, a critical remark on subjunctive conditionals (the second premiss of abduction has this logical form) and counterfactuals opens the door to a reflection about the plausibility or implausibility of information contained in them; and on the other hand, the contributions of Ronald Langacker's *Cognitive Grammar* can help to clarify the role played by accessing and activating certain zones of conceptual structure in determining the relevance of information to agendas-in-course.

The paper ends with some open questions about the need, if any, for explicit commitments for the project of naturalizing logic.

2. Practical logic and cognitive systems

"Accordingly, we will say that a cognitive system is a triple of a cognitive agent, cognitive resources, and cognitive target performed in real time" (2009, p.628, Quantum Logic).

Agendas and Agents

Agentive cognitive work: right response versus correct response versus truth

I agree with Woods that there is a conceptual priority of reasoners over reasoning (Woods, 2007, p.75).

I will say here that an agent is an entity with intrinsic or operative agency, that is, capable of acting; and the aim he/she is trying to achieve constitutes his/her *agenda*². According to Woods, a practical individual agent has some cognitive resources: "information, time, storage and retrieval capacity, and computational complexity" (Woods, 2007, p.77), plus modes and methods of inquiry (Woods, 2013, p.72). I would like to add not only cognitive functions such as attention, working memory, and executive function, but also other dimensions of agency such as sensory-motor capacity, affectivity, sense of space and time, and intersubjectivity. That is, an agent is embodied/animated. Besides, an agent engages (and has transactions with) his/her world, in a multilayered set of practices, and in doing so, he/she situates himself/herself in it. That is, an agent is situated.

The agenda an agent is aiming at is a type (in the technical sense of something that comprises variation) of outcome for which something is selected (Short 2007, p. 92–93). By this definition, agendas include –as in Woods (2013, p.201)– plans, intentions, strategies, and inclinations³; and they can range from completely "dark" (unconscious,

² This section is heavily based on Niño & Marrero (2015), which in turn is based on Niño (2015).

³ Woods includes things like functions, programs, scripts, and matrices. In my present view, machines and artifacts are not agents, since their agency is not intrinsic, but derived or assigned. Strictly speaking, one does not interact with a computer, because a computer does not 'act', in this sense of agency.

automatic, inattentive, non-voluntary, non-conceptual) to completely ‘light’ (conscious, deliberate, attentive, voluntary, conceptually structured) (cf. Woods, 2013, p.201).

I will use the term “agentive” for anything related to agents, and “agendive” for anything related to agendas.

When an agent is trying to fulfill his/her agendas, we have to keep in mind two kinds of *resolution conditions* for the agenda: the final state or situation that is aimed at –or *closure conditions*; and the *scope conditions*, or the more salient or probable or typical resources the agent expects to use or is using in order to obtain closure. For instance, if the agenda is to have ramen for dinner, the closure conditions involve having had ramen for dinner, and the scope conditions include the way that state is obtained: For instance, you can prepare the ramen, ask a friend to prepare it, go to a Japanese restaurant, order it for home delivery, etc. Notice that closure conditions are determinable independently of actually eating the ramen: it has to be ramen and not soba or another food, and it has to be for dinner and not lunch, or brunch. In this sense, closure conditions are conditions to which the agent subjects himself/herself in order to accomplish his/her agenda (Woods, in *Seductions and Shortcuts*, agrees with this). Probably, an agent is free to choose (most) of his/her agendas, but not their closure conditions. In relation to scope conditions, which ones are really available (or at least, experienced as available, viable, and possibly executable) depends on the agent’s resources. As a whole, resolution conditions function as a reference point in order to evaluate an agent’s real performance.

Agendas’ embeddedness

Now, the agent’s actual execution of actions, in order to fulfill his/her agenda, is the *agentive performance*. For instance, in the above example, if the agent decides to prepare the ramen himself/herself, the agentive performance includes the actions for cooking the ramen, obtaining the ingredients, etc., and, of course, the intake. If the agent meets the resolution conditions, he/she then ‘closes’ his/her agenda.

The constraints that resolution conditions impose on the agent establish a certain *standard of precision* for their fulfillment. They include the degree of accuracy, fluency, and exigency with which the agent expects⁴ to close his/her own agenda (Niño 2015, p. 108). For instance, in preparing the ramen, the agent could expect a certain degree of freshness in the ingredients and a certain quantity for the final meal, and his/her actions are directed toward meeting those expectations. These standards of precision contrast sharply with the standards socially imposed on the agendas (presumably) attributed, assigned, or imposed on an agent, because of the social role he is supposed to fulfill. And this is so because what counts as “good enough” or “not good enough” for the agent does not necessarily coincide with what is socially expected from someone in his/her role. This is why one expects greater accuracy, fluency, and exigency in the performance of a

⁴ “Expects” is more appropriate for light than for dark agendas. But dark agendas also have some standard of precision: most of the time we are too inattentive to notice our breathing, but if we start to breathe with difficulty, we notice. In the nineteenth century, Leriche said that health is the silence of the organs. As a definition of health, it is not good enough, but it allows us to see that an organism’s keeping itself alive, when no one cares about it, is a dark agenda.

professional pianist, as opposed to the performance of an amateur. I refer to these socially constructed standards as *standards of strictness* (Niño 2015, p. 128). An agent can align or not align his/her own standard of precision with the socially imposed standard of strictness (and this alignment/misalignment can be produced in light, in darkness, or by chance). When performance is evaluated in relation to the standard of precision, the result is called the *degree of precision*. This refers to the level of specificity to which an agent advances his/her own agendas (Niño, 2015, p. 131). On the other hand, if performance is evaluated in relation to the standard of strictness, the result is called the *degree of strictness*, and refers to the level of the performance according to the social expectations (Niño 2015, p. 128).

Different cultures and societies, in different periods of history, establish different kinds of standards of strictness for clusters of agendas associated with social roles, which I call *agendive roles* (Niño, 2015, p. 125–126). Examples of agendive roles are cop, surgeon, engineer, cellist, astronaut, prosecutor, soccer player, friend and parent. Each role is composed of different activities, which have their own purposes. For instance, teachers are expected to follow the syllabus in order to achieve the goals of the course, to prepare their classes in order to facilitate learning, to prepare exams in order to evaluate the level of learning of their students, etc. Boxers, in turn, are expected to train before the fight, do strength exercises, have a strategy to test with a sparring partner, etc. Now, the same activity can be present in different roles: parents and cops interrogate; parents and singles cook, people have conversations, etc. In turn, activities can be understood as clusters of actions (which have their own aims). And, of course, one action can be present in different activities: both politicians and partners make promises. So, we have standards of strictness for agendive roles, agendive activities, and agendive actions. Not always, but very often, the standard of strictness for an action depends on the activity it is embedded in: The determination of a certain body temperature as fever would require different standards of strictness depending on whether the agent is Joe's father or Joe's pediatrician.

Actions can be divided into kinetic (locomotive, manipulative, and exploratory), expressive (including speech acts), and comprehensive. The latter include acts involved in making inferences, imagining the possible and impossible, perceiving, remembering, calculating, etc. (Niño, 2015, p. 111-114). Of course, in one activity, several different kinds of actions can be involved: in cooking, it is important to manipulate the ingredients, taste them, calculate portions, perceive different states of the preparation or even read and follow instructions, etc. Insofar as actions can be socially assessed or internally evaluated, both standards of strictness and standards of precision apply to these three kinds of actions.

Now, because of their intersubjective nature, agendive roles behave in tandem: we have doctors, patients, nurses and paramedics; judges, juries, defendants and prosecutors; professors, teaching assistants, monitors and students; parents, siblings, sons, daughters, etc. I understand *context* as the sociohistorical construct in which agendive roles intersect (Niño 2015, p. 146–147), and *topic* as the interlacement of purposes and/or activities. For instance, one topic of the educational context where the roles of student and teacher interact would be the interest in learning (Niño 2015, p. 150–151). In the case of patient

and doctor, one topic would be the interest in the preservation of health or recovering from a disease. When an agent accepts the purposes of a role, enacts the corresponding activities or takes on the responsibility for carrying out an agentive role, he/she endorses an *agentive role* (Niño 2015, p. 146–147). As actions are embedded in activities, which in turn are embedded in roles, we also have agentive actions and agentive activities (of course, not all actions are embedded in social roles). As before, agentive roles are evaluated against the corresponding agentive roles; and thus, by a certain cluster of standards of strictness. Therefore, the activities an agent performs should be assessed within the framework of the activities attributed to his/her agentive role. For instance, it is not expected that the semiological competence of a medical student surpasses that of a medical doctor. Finally, the topic in which the performance is embedded can change depending on the role the performer is playing. To illustrate, the topic changes if the medical student is interacting with his/her patient, or with his/her professor of medical semiotics.

But this being so, we can see that a *purely epistemic agenda* (for a practical agent) is something difficult to find. Normally, epistemic agendas are at the service of other kinds of agendas; that is, the epistemic dimensions of a practical agenda include comprehensive actions, whose purposes are embedded in non-epistemic agendas. A medical doctor tries to determine a diagnosis as a means for therapy. A prosecutor examines and weighs evidence to establish guilt in a crime. The contestant in ‘Who Wants to Be a Millionaire?’ needs veridical information in order to win the prize. In contradistinction to the search for fun, emotion or food, the search for truth for its own sake (praiseworthy as it is) is not very common. And it is too expensive for a practical agent.

Keeping this in mind, I want to introduce two additional ways to understand agendas. First, in relation to the *agency* involved: (a) if the resolution conditions require only the agency of the agent who advances the agenda or (b) if they require the participation or modification of another agent’s (or agents’) agency. Second, in relation to the *action’s effects*: (i) if the agenda closes at the very moment the actions are performed or (ii) if it closes through the subsequent effects of the actions. These criteria intersect with each other, producing four categories. However, I want to discuss here just (ai) and (bii). Case (ai) occurs when an agent acts and closes his/her agenda at the very moment the action ends, as a result of that action rather than, for instance, by chance. For instance, I want to go to the second floor from the basement. I climb the stairs, and at the moment I step from the last step to the second floor, I close my agenda. (ai) is the usual way Woods presents abduction (cf. Woods, 2013, Ch. 11). Case (bii) occurs when an agent acts, and because of the effects of his/her intended actions, another agent acts, and with this action of the second agent, the first agent closes his/her agenda. Communication is a paramount example: Peter asks Mary Jane to raise her arm. He produces a series of sounds with his mouth and makes a speech act. He calls Mary Jane’s attention, she understands what he says, and, as an effect of her understanding, she raises her arm, and then, Peter closes his agenda (I call those agendas *per-agendas*, from the Latin “*per*” which means “action carried out until its end”, cf. Niño, 2015, p.120-121). Of course, she could have refused to raise her arm. There is a gap between agencies sometimes called ‘freedom’. (bii) is typical

of being told about things (cf. Woods, 2013, Ch. 9), but also of the game of giving and asking for reasons in arguing. Thus, argumentation, as an activity, is a kind of per-agenda.

Let's look at the latter closely. Arguing is an activity associated with different agendive roles, but it is not a social role by itself: there are no "argumeters" as there are teachers or lawyers, even if teachers and lawyers argue. Consequently, the standard and degree of strictness of arguing are relative to the role or roles to which it belongs. This is why arguing is very different for a mathematician, a medical doctor, a broker or a merchant. In this sense, what constitutes good or bad argumentation (or even, good or bad reasoning in argumentation) involves the kind of agendive role and standard of strictness in which this or that arguing is embedded. But we must keep in mind that just as there are standards of strictness, there are also standards of precision; and the agent, in advancing his/her agenda, can align or misalign these standards of precision with those standards of strictness. Perhaps he/she is too demanding or too relaxed for the given circumstances. What distinguishes argumentation from other kinds of activities, such as being amazed by a magic trick or entertained by a story, is that argumentation causes the modification of an agent's agency as an effect of the arguments, and not as an effect of other things such as threats, spite, drugs, etc. (cf. Niño & Marrero, 2015).

Now, even if *purely* epistemic agendas are rare, I want to propose four kinds of agendas in which the activity of arguing can intervene. They are not intended as a complete list, but as an example of the uberty of the approach⁵. I will relate "epistemic" to the standard of strictness and "doxastic" to the standard of precision. In this sense we can find: agendas of epistemic/doxastic (1) arrival; (2) maintenance; (3) defense (and presentation); and (4) obstruction and criticism.

In the first, agendas of epistemic/doxastic arrival, the agent's aim is to resolve ignorance problems (see below) that are irritating him/her, and thus, favor the formation of opinion. This class includes cases of looking for information in a book to clarify a question or by asking someone. But this class also includes the verification of a scientific hypothesis. In advancing a scientific hypothesis (see below), an agent gives reasons to someone or to himself/herself for accepting, or at least, for taking into consideration, the possible explanation put forward. In addition, it also covers cases in which, in the absence of appropriate and complete information, some information (even if rough or merely approximate) would suffice to make a decision: sometimes, in an urgent dilemma, an educated guess could be enough. Of course, the accuracy and precision of the information required or admissible depend on the level of strictness or precision. Homework about "eukaryotic cells" is not usually the same for a high school student or a PhD student.

In the second class, agendas of epistemic/doxastic maintenance, the aim of the agent is to preserve or reinforce an already formed opinion. This class includes cases where people already share an opinion about something, and the per-agenda consists in the advancement of reasons in order to root and deepen that opinion in the other agent's

⁵ A previous version of three of these four classes of agendas was proposed in Niño (2013, pp. 233–236); and Niño & Marrero (2015).

mind. For example, if two people already agree that Batman is better than Superman as a hero, their mutual giving and asking for reasons to favor Batman's superiority over Superman deepens their previous opinion.

In the third class, agendas of epistemic/doxastic defensibility, the agent's aim is to defend a previously formed belief (fixed by the first class at some prior moment). Take, for instance, a debate between politicians of different and opposite political parties about corruption. Each one is advancing reasons about previously formed opinions on the origin of corruption. Note, however, that the target of those reasons is not the arguer himself/herself, but another agent: a potential voter. In this sense, persuasion is a paradigmatic case in agendas of defensibility. Here the per-agenda consists in gaining someone else's assent to an opinion. Persuasion is a very heterogeneous phenomenon. This is so because the modification of the target's agency can be easy or difficult, varying along a wide range of admissibility conditions for multiple audiences, as well as a very diverse spectrum of standards and degrees of strictness. For instance, it is not the same to defend the benefits of vaccines before doctors who prescribe vaccines of another brand, before an anti-vaccine audience, before a layperson who has never heard of vaccines, or before potential funders of further research. In this sense, the resolution conditions of this class of agenda take into account the admissibility conditions for the persuasible target (which can vary along a spectrum: the same argument can be shown to different audiences with different attitudes towards the argument and the arguer). These conditions include the verisimilitude of the facts and the acceptability of values involved in the construction of the reasons, as well as those associated with the arguer (see below). For instance, in empirical science or mathematics, where the standard of strictness in methodology is (presumably) high, and thus, the facts established are filtered, an arguer, based on this, can advance his/her reasons, even if the facts are not very plausible. Think, for instance, of the impossibility proof for squaring the circle. Many generations of geometers thought that it was possible, but it is not. In a certain sense, that was shocking, but irrefragable. In science, presumably, the acceptability imposed by the methodological constraints overrides everything else. It is very different if the discussion is about politics. Here, both acceptability and verisimilitude can count, if not equally, at least in a different relation from that found in empirical science or mathematics. Therefore, the audiences' epistemic/doxastic system becomes crucial for the achievability of the agenda. Is the audience hostile or sympathetic to the topic or the arguer? Is it liberal or conservative? Is the topic controversial (e.g. pro-choice versus pro-life) or not?

In the fourth class, agendas of epistemic/doxastic obstruction and criticism, the agent's goal is to attack a previously formed opinion (fixed by the first class at some prior moment) of his/her target (criticism) or to prevent him/her from forming an opinion or presenting it to a further audience (obstruction). In this case an agent A attacks agent B's reasons for thinking that a is the case. In this sense, criticism tries to put the admissibility system against the agent himself/herself: his/her reasons for acceptability/verisimilitude are flawed: So, Joe in talking to Jane, says: "You cannot believe that Johnny is faithful, when it has been proven that he is a hopeless cheater!". But, maybe Joe does not only want Jane to stop believing in Johnny's innocence. Perhaps he wants to prevent her from

forming new beliefs about Johnny: he also mentions other reproachful behaviors of Johnny, the bad company he keeps, etc.; and in doing so, Joe obstructs Jane's agenda. (Maybe, even, he distracts her, by saturating her with more information than necessary, making her waste time, etc.).

Of course, these four classes of agendas can overlap in a given activity of arguing. But in no case is arguing the ultimate or final agenda. It is always a means to another, non-argumentative end.

I think the distinction between standards of precision and standards of strictness reveals something else about some of Woods's theses. For instance, Proposition 1.8d. *Errors of omission* says "Some errors of reasoning are consequences overlooked. In other words, the consequence recognition task is sometimes not successfully performed" (Woods, 2013, p.26). Now, suppose SG is a medical doctor in an emergency room. She is very demanding of herself. She overlooked a diagnosis, but nobody blames her: it was a very rare disease and she could not be expected to rule out other diagnoses and arrive at that one in such a short time. However, she feels guilty: she did not draw the diagnosis from the symptoms, or at least, did not think of a test in order to determine that diagnosis. So, she falls short of her own degree of precision but exceeds her degree of strictness. We may think she is too severe with herself. Now, think of the following case: GS is a mediocre student admitted to a high-quality school. He is content with mediocre results, notwithstanding his other privileges. At the end of the term there is an exam. He gets a barely passing grade. Here GS exceeds his degree of precision but falls short of his degree of strictness.

One more example: according to Peirce, in a situation of scientific research, we have the right (and duty) to entertain the abductive conclusion as a suspicion (an interrogation or a question, even), but not as a belief (see below). In scientific research belief is authorized by the inductive conclusion. Now, suppose GS arrives at a scientific hypothesis, and at that very moment, he decides to believe it, instead of only "suspecting" that it would be true. If Peirce is right, GS makes a mistake: he accepts as actual what is merely admissible as possible. Now, again, in a very urgent situation (that is to say, with very little time), like in an emergency room, we have the right to act immediately on the information the hypothesis affords, because otherwise, the main agenda becomes unclosable (waiting too long in an emergency room could be –literally– a fatal error). Now, suppose SG arrives at a diagnostic hypothesis in an emergency room, and without waiting a second acts on it. If Peirce is right, SG does not make a mistake, even if her diagnosis is incorrect (see next section).

This has implications for the NN Convergence Thesis and, *a fortiori*, for the project of naturalizing logic: the NN Convergence Thesis holds when both standards are aligned, but problems for the thesis can arise when they are misaligned.

[Working note: NN and attributed agendas.] The resolution conditions are expected to conform to certain standards that a culture, at a given time, establishes for their fulfillment.

[Working note.] Efficient cause. Agenda: final cause (what devices “are meant to”). What do we do with that? How do we naturalize that? It is plain that people have purposes. Sober!

[Working note.] CRM and per-agendas: it is more natural to say that we believe what we are told because that is the expected effect assertion has.

3. Abducting abductions: from hopeless despair to educated guessing

If the naturalization of logic seeks to understand how real reasoners reason, we should look for what those ways are. Insofar as those ways meet neither traditional deductive nor traditional inductive standards, Woods calls them “third-way reasoning”. He even adds that finding third-way standards is the “more important and widely open problem for the logic of error” (2013, p.67). There are various examples of third-way reasoning: hearsay, hasty generalization, and abduction. So, abduction is a paradigmatic form of third-way reasoning and Woods has devoted several influential works to it. Woods usually begins his discussion on abduction with what is known as the *Peircean abduction schema* (Peirce, 1903, p.231):

The surprising fact *C* is observed.

But if *A* were true, *C* would be a matter of course.

Hence there is reason to suspect that *A* is true (CP 5.189, 1903).

Woods (2013, p.365) highlights that, in his opinion, Peirce’s abduction is characterized by the traits of surprise, suspicion and matter of course, which he calls *the Peircean Triad for abduction*. (I hasten to say here that this triad is directly related to the *conditions of admissibility* I mentioned in the previous section, and that I will return to them in this and in the next section, taking into account what was said there).

Then, he makes seven comments about the way he understands Peirce’s abduction (Woods, 2013, p.365):

- P1. Abduction is triggered by surprise [CP 5.189].
- P2. Abduction is a form of guessing, underwritten innately by instinct [CP 5.171, Peirce, 1992, p.128].
- P3. A successful abduction provides no grounds for believing the abducted proposition to be true [Peirce, 1992, p. 178].
- P4. Rather than believing them, the proper thing to do with abducted hypotheses is to send them off to experimental trial [CP 5.599], [CP6.469–6.473], [CP 7.202–219].
- P5. The connection between the truth of the abducted hypothesis *A* and the observed fact *C* is subjunctive [CP 5.189].
- P6. The inference that the abduction licenses is not to the proposition *A*, but rather that *A*’s truth is something that might plausibly be suspected [CP 5.189].
- P7. The “hence” of the Peircean conclusion is ventured defeasibly [CP 5.189]).

In (2017, p.138), Woods adds: “Let us note that P3 conveys something of basic importance. It is that successful abductions are evidentially inert. They offer no grounds for believing the hypotheses abducted. What, then, is the good of them?”. And from that, he advances the idea that when facing *ignorance problems* an agent can respond in three ways: (a) subduance, if the new knowledge removes ignorance and that new position serves as a positive basis for new action. With subduance, the agent overcomes his/her ignorance; (b) surrender, if ignorance is fully preserved in a way that this state cannot serve as a positive basis for new action. With surrender, ignorance overcomes the agent. Or (c) abduction, which like surrender is ignorance-preserving, but like subduance, offers a basis for new action. With abduction, ignorance remains, but the agent is not overcome by it (Woods, 2017, p.139). As we saw above, arrival agendas cover these three cases, with full closure in subduance, full abandonment in surrender, and expectation of advance without closure in abduction.

Once Woods offers the epistemic framework in which he understands the role of abduction, he proceeds to show what is known as the Gabbay-Woods Schema of abduction (of which there are various versions, cf. 2005, p.47; 2013, p.369; 2017, p.139). Let me quote Woods *in extenso*:

“let T be an agent’s epistemic target at a time, and K his knowledge base at that time. Let K^* be an immediate successor of K that lies within the agent’s means to produce in a timely way. Let R be an attainment relation for T and let $\rightsquigarrow$ denote the subjunctive conditional relation. $K(H)$ is the revision of K upon the addition of H . $C(H)$ denotes the conjecture of H and H^c its activation⁶. Accordingly, the general structure of abduction can be captured by what has come to be known as the Gabbay–Woods schema:

1. $T! E$ [The $!$ operator sets T as an epistemic target with respect to some state of affairs E]
2. $\neg R(K, T)$ [fact]
3. Subduance is not presently an option [fact]⁷
4. Surrender is not presently an option [fact]
5. $H \notin K$ [fact]
6. $H \notin K^*$ [fact]
7. $\neg R(H, T)$ [fact]
8. $\neg R(K(H), T)$ [fact]
9. $H \rightsquigarrow R(K(H), T)$ [fact]
10. H meets further conditions $S_1, \dots S_n$ [fact]
11. Therefore, $C(H)$ [sub-conclusion, 1–7]
12. Therefore, H^c [conclusion, 1–8].

⁶ In Woods, 2011, he adds: “Activated conjectures are “working hypotheses”” (p.241).

⁷ Steps 3 and 4 are new in relation to the presentation of Woods (2013, p.369). There, however, the following was presented as step 3: “ $\neg R(K^*, T)$ [fact]”, which in this presentation is missing. I think this is so, insofar as K^* , as a successor of K , does not overcome ignorance sufficiently to close $T!$

It is easy to see that the distinctive epistemic feature of abduction is captured by the schema. It is a given that H is not in the agent's knowledge set K . Nor is it in its immediate successor K^* . Since H is not in K , then the revision of K by H is not a knowledge successor set to K . Even so, $H \rightsquigarrow R(K(H), T)$. But that subjunctive fact is evidentially inert with respect to H . So the abduction of H leaves the agent no closer than he was before to achieving the knowledge he sought. Though abductively successful, H doesn't enable the abducer to attain his epistemic target. So we have it that successful abduction is ignorance-preserving. Of course, the devil is in the details. Specifying the S_i is perhaps the hardest open problem for abductive logic" (Woods, 2017, p.139).

In this section I want to make some general remarks concerning the Woodsian approach to abduction and the G-W Schema. In the next section I will make some remarks in relation to the challenge posed by the 'hardest problem' of S_i .

First of all, I want to say I sympathize with the general approach Woods takes to abduction. I prefer it over others. Now, anyone who studies abduction ends up mentioning the famous Peircean abduction schema. (For some people this is canon, for others, not so much). Woods fortunately takes the trouble to track down seven insightful Peircean ideas on abduction: P1-P7. I think Woods develops all of them very well. But I think P4 deserves two comments. Let me quote it again:

P4. "Rather than believing them, the proper thing to do with abducted hypotheses is to send them off to experimental trial (CP 5.599; 6.469-6.473, 7.202-219)" (Woods, 2013, p.365).

Well, for certain kinds of abduction, but not for all of them. And I'm not thinking of the "rather than believing" part, but of "the proper thing to do" part. And for this I will quote Peirce:

the logical rules of inference from a surprising fact to a hypothetical state of things that would explain it involve a complexity of conditions, one of which is how much time can be allowed for coming to one's conclusion. A general who during a battle must instantly risk the existence of a nation either upon the truth of a certain hypothesis or else upon its falsity, must perforce go upon his judgment at the moment; and his doing so is in so far logical that all reasoning is based upon a tacit assumption that Nature, in the sense of the aggregate of truth, is conformed, more or less, to something similar to the reasoner's Reason. This kind of inference may be called Practical Retroduction. It is ordinarily called a Presumption. If on the contrary, indefinite time can be allowed for judging of the truth of an explanatory hypothesis, the most favorable conclusion is that the consequences of that hypothesis that are susceptible of verification should be subjected to systematic and thorough tests. This, which may be called Scientific Retroduction, will be subject to considerations of economy. If, for example, a supposition far from reasonable, yet still possible, would explain a physical phenomenon, and if, in case it be incorrect, there is a way of disproving it at little expenditure of either time, energy, or any other valuable, then it may be worthwhile to clear the ground by taking up the investigation of it forthwith. It will be remarked that the result of both Practical and Scientific Retroduction is to recommend a course of action" (MS 637, ISP4-6, 1909).

Now, this is interesting. Peirce's distinction between Practical and Scientific Abduction (here called Retroduction⁸) is related to time, understood precisely as a resource for the

⁸ Peirce called the inference to an antecedent from a consequent and a consequence, "reasoning *a posteriori*", "Hypothesis", "Presumption", "Abduction", "Retroduction", "Reasoning from signs", etc. "Abduction" was used in 1900-1906; "Retroduction" in 1898, and from 1906 until one month before his death, in 1914. This means, by the way, that "abduction" was not his final choice for that form of inference.

resolution of problems⁹. Note that the question of time has an impact on the *severity* of the result (or the urgency with which the response is required) of the hypothesis under consideration, and not on the *type* of problems that are being solved. Therefore, the distinction between Practical and Scientific Retroduction is not that of greater or lesser conceptual difficulty, but that between what has to be resolved *urgently* and what can be resolved without that imperative. The distinction that is established here, then, is the distinction between a type of activity that requires the solution of problems whose solution cannot wait and a type that permits delay. So, “the proper thing to do” with an abductive conclusion depends on the kind of abductive situation one is in. Of course, one can imagine a large range of intermediate cases between Practical and Scientific abductions. I mentioned above the case of a medical doctor in an emergency room¹⁰. She can’t wait too long, because if she does, her patient can die. This case is more practical than “scientific”.

Another thing to note in the example for the Practical Abduction is that what is offered is not an explanation but a solution to a problem, and of course, an unwarranted solution. So the lesson seems to be: better a possible solution (even if false and destined for failure) than none at all. Woods says something similar about abduction for individual practical agents: “It answers to the idea that there are situations in which it is better to do something than nothing.” (2011, p.242).

Now, concerning Scientific abduction, Peirce recommends some “Economy of research” (distinct from cognitive economics). Under the item “Non Complexity” Peirce says:

There still remains one more economic consideration in reference to a hypothesis; namely, that it may give a good “leave,” as the billiard-players say. If it does not suit the facts, still the comparison with the facts may be instructive with reference to the next hypothesis. For example, I might be inclined to surmise that an observable quantity y was such a function of a quantity x , determined by the conditions of experiments, as to be expressible in the form $y = a + bx^2$. Still, as I am not sure of this, perhaps it would be wise first to try how well the experiments could be satisfied by $y = cx$; because the residuals will be more readily interpretable in the latter case. As a provisional hypothesis, it will, for this reason, other considerations apart, be better to assume something very simple, even although we imagine that by complicating the hypothesis it could be brought nearer the truth (CP 7.221, 1901)

Although the hypothesis does not give an account of the facts in a completely satisfactory way (and so, is lacking some of the “matter-of-courseness”), it can still be illuminating for the construction of the following hypotheses. Thus, it is preferable first to adopt conceptually simple and even rough hypotheses, even knowing, in advance, that they are false (CP 7.222). This is compatible, besides, with what Peirce considers the first “fallacy of abduction”: “The commonest fallacy of retroduction [i.e. abduction], as it seems to me, is

The chronology and explanation of all those uses can be found in Niño (2008).

⁹ A great deal of time and cognitive resources may be expended between the “surprising fact” of the first premiss and the “matter of course” of the second. Inquiry begins with the surprising fact, not with the “matter of course” (The Eureka Moment).

¹⁰ I would add Peirce’s biographical anecdote of the stolen watch (1907, MS 687: ISP 9-24) –made famous by the Sebeoks– as a case in which, even without any conscious basis, the conclusion of the guessing process impels action.

the idea that the most probable hypothesis is the best... A much more useful maxim is, be upon your guard against assuming anything to be true because it seems likely or a matter of course; for that is the great source of delusions" (RLT, 1898, p.193). Therefore, for Peirce the best hypothesis for scientific research is not the most likely (not even the 'loveliest' as inference to the best explanation theoreticians would like), but that which can be conveniently, economically, and quickly put to the test in order to advance in scientific inquiry.

Thus, there are cases in which the hypothesis recommended is advanced, not because it is presumptively true, but because it is presumptively good for the advancement of inquiry. But of course, as a rule of thumb, a presumptively true hypothesis is, presumptively, good for the advancement of scientific inquiry.

So, if –on Peircean grounds– truth is the ultimate goal of scientific inquiry, the Peircean abduction schema captures that, which means two additional things: first, the Peircean schema is *typical*, but not universal (nobody thinks it is), even for his own idea of scientific research. Second, what is after the "hence" may not be true, but it may be valuable enough to justify investing resources in its investigation. I think the Peircean abduction schema does not need any amendment. I just want to remark that the schema –and particularly, its abductive conclusion ("Hence there is reason to suspect that A is true.")– was designed for scientific research. But, insofar as scientific research is a sophistication of our everyday practices, what is typically true of scientific abduction is a projection of our "fast and frugal" practical abduction.

What I see here is that Peirce's distinction between practical and scientific abduction intersects Gabbay and Woods' distinction between practical and institutional agents. Peirce's practical and scientific abductions can be advanced by individual agents (at least, for him). Peirce's examples of scientific abductions include Kepler's discovery of the orbit of Mars and Dalton's atomic theory; that is, individual agents, but scientific men. And of course, one can imagine some NASA research programs beginning with scientific abductions. But crucially, both distinctions have *time* as a distinctive criterion. The scarcity of time can mean very different things for an individual practical agent. If an agent is playing in a guessing game (say, Pictionary), and has one minute to guess what his/her companion is clumsily cueing, he/she is in a practical abductive situation, but not like that of the emergency room or the general in the middle of a battle. One can even suppose that the guessing range can be extended or narrowed depending on what is at stake: maybe the whole wide range between wild and educated guesses is at some point admissible in the game, but only educated guesses are admissible in an emergency room¹¹. Maybe something analogous could be said about the range of "matter-of-courseness" the range of guesses can afford. And this leads me to another comment: what

¹¹ Let's put it this way: there are two extremes in guessing. On the one hand, the wildest guess in which no reason whatsoever can be given to point in the direction pointed out, as in Peirce's example of the investigator who treats the first name found randomly in an address book as that of a murder suspect (this, *prima facie*, seems gratuitous and unreasonable). On the other hand, the educated and accurate guess, which leaves little room for thinking that bad reasons are backing up the conclusion.

is admissible or not, will depend on the resolution conditions of the agenda-in-course for the agent (standard of precision) and/or the resolution conditions of the attributed/assigned/accepted agendas for the agendive action/activity/role (more about the “matter-of-courseness” is said below in the item ‘admissibility’).

Now, in my view P2, P3, P5, and P7 are not touched by the distinction practical/scientific abduction. It is pretty clear that we are not allowed to believe (P3) our guesses (P2) in the sense in which, for Peirce a belief is a habit of action. For Peirce, if someone says he/she believes that p , he/she has positive expectations about what to perceive, what to do, or what to think about the objects referred to in the scope of p . But this means that that agent is non-doubtful about those expectations. That is why if those expectations are disappointed, a surprise emerges. Think of this as a bet. If you are pretty sure that the next card is an ace of clubs (suppose, you can count cards), you form a positive expectation about it, and you are confident in making a bet based on it. Now the next card is a nine of hearts, and you are surprised (and frustrated). But now suppose you have no clue about what the next card is in the first place. But you are urged to bet because those are the rules of the game and it is your turn. You have a hunch that the next card is an ace of clubs but it is a nine of hearts. You are disappointed (you have just lost a substantial amount of money), but not surprised. There was no warrant that you were right. Now a third scenario: You have no clue what the next card might be, but you bet on the nine of hearts. And the next card is a nine of hearts. You are glad and surprised. You don’t say “I knew it!”, but something like “That’s funny. What were the odds of winning?” So the level of surprise is related to the level of confidence you put in the information. Sometimes, however, you lack confidence in the current information, but are urged to act on it. This, on the other hand, is compatible with not-knowing in the CRM of knowing: The agent does not believe a (Woods, 2013, p.93) and the information is not good because it is not complete, even if accurate (Woods, 2013, p.114). The absence of permission to form a belief is not an obstacle to acting, but it is a criterion of evaluation of epistemic behavior: if an agent believes his/her abductive conclusions, he/she behaves poorly epistemically.

This leads to another comment, this time on P6: In scientific abduction we have the permission to suspect the abductive conclusion is true. And perhaps also in some cases of practical abduction, like the doctor in an emergency room. But it is possible to think of cases where there is no ‘suspicion’, but a ‘hope’ that the abductive conclusion be true (or be a solution or be a worthwhile investment). This is the situation of the general (or, as I will call it, Captain America’s hopeless hope in Avengers: Endgame: “It will work because I don’t know what I will do if it doesn’t”).

I would add, concerning the G-W Schema, that, if the distinction practical/scientific abduction is worth taking into account, it would have some bearing on line 1 (the type of target) and line 10 (because further considerations are different for each one).

Now, we come to the *Ignorance-Preservation thesis*. Under the criteria of the CRM abduction does not produce knowledge. There is room neither for belief nor for good information. If it results in belief formation the agent’s epistemic conduct is questionable (it is too much to say that something is wrong with the belief-forming devices without

empirical research, but I suspect there are cases of this sort, like in pathological jealousy). If information is good, it is by chance. Everything after *K* should be considered as epistemically sub-par (Gabbay & Woods, 2006). By the Error Abundance Thesis (Woods, 2013), we think that errors in reasoning due to misinformation outnumber errors of reasoning. In this case, then, the error of reasoning is to believe the abductive conclusion. And it is easy, on Peircean grounds, to see why: if you believe the abductive conclusion, you halt scientific inquiry at its first stage: if you already believe that *p*, and have no doubts about it, why should you continue to determine its truth? As far as I can see, there is nothing in P1-P7 which can override all of this. According to Woods, the goal of abduction is “to make the best of the ignorance that one chances to be in.” (2017, p.139).

However, Lorenzo Magnani is not convinced. Magnani says Peirce himself was willing to say that perception is abductive and quotes him: “Abductive inference shades into perceptual judgment without any sharp line of demarcation between them” (Peirce, 1903, *apud*, 2018, p.10). This implies that P1 has to be rejected as a feature of abduction, because the majority of our everyday perceptions are unsurprising (unless, “surprise” means something different). And it implies that P3 should be rejected, because our perception contents become our online (and offline, via memory) beliefs. By doing this, Magnani throws away two features of Woods’s Peircean Triad for abduction (see above). Now, I am unconvinced. Peirce relates abduction and perception occasionally. The following is one such occasion:

Now, resume the division of reasoning. You begin with *perception* which brings *surprise*. To make that reasonable you resort to Retroduction. This resembles perception in *bringing* the *new*. And after that no new idea enters into reasoning (MS 754: ISP 5, 1907)

What triggers abduction is something curious, unusual, that demands a solution (be it an explanation, a strategy, etc.). Perception functions as an information-capture strategy in its own right. Perception and abduction pertain to that kind of thought movement that brings something new: online information in the first case, new ideas in the second. But they are very different. It is true, Peirce relates perception and abduction, but also distinguishes them. Of course, Peirce can be the first authority to consult on abduction, but he cannot be the last. Why should he be? However, Magnani does not tell us what to do with P1 or P3. Are they to be replaced? If we include perception-as-abduction as a kind of abductive inference in contrast with ‘sentential’ abduction, captured in the G-W schema, what do we gain? (cf. Woods, 2011, p.243).

P5 is related to the subjunctive character of the second premiss of the Peircean abduction schema and P6 with the plausibility (for suspecting) of its conclusion. Those are issues that will be examined in the following section.

4. Information: relevance and plausibility

Woods gives several reasons for the idea that *relevance* and *plausibility* (and their counterparts, irrelevance and implausibility) are very important notions for a naturalized logic (cf. 2013, p.28-29). Relevance and plausibility, however, are predicated of information. According to Woods (2018, p. 59-60), information has four senses: epistemic,

probability, complexity and military. This ambiguity makes it very fruitful. In this section I want to discuss some properties of the epistemic and complexity senses, which can have a direct bearing on plausibility and relevance.

Information, conditionals and drawing consequences

The second premiss of abduction (“but if *A* were true, *C* would be a matter of course”) presents many and very difficult problems: How is *A* selected over other candidates? How is *A* generated (if new)? What does it mean for *C* to follow from *A* as a matter of course? The latter has to do with conclusionality, the second with the subjunctive character of the premiss. Both of them, I think, are related to counterfactual thinking¹².

It is usually supposed that the typical way of expressing a counterfactual is by means of a subjunctive conditional (cf. Bennett, 2003, for objections to this denomination). I will not object to this here. I will focus on the problem of what are known as “Anderson-type examples”. As is well known, these ‘examples’ arise by contrasting a conditional in its indicative with its subjunctive mode of expression. Let’s begin with a non-problematic example:

(1) John’s waking up

(1a) If John did not wake up at 7:00am, he arrived late for the meeting.

(1b) If John had not woken up at 7:00am, he would have arrived late for the meeting.

This seems pretty straightforward. The difference between the indicative and the subjunctive is that in the indicative if something did not happen actually, then a second thing actually happened; while in the subjunctive, if something had not happened, then a second thing would have happened. And the first and the second thing seem very similar, the differences being in the mood/aspect. But, as it has been noted for decades, things are not always so easy. Sometimes the conversions are less smooth:

(2) Oswald (from Adams, 1975):

(2a) If Oswald didn’t shoot Kennedy, someone else did.

¹² Why do people make counterfactuals? Ruth Byrne (2016, p.1) says that “Counterfactuals explain the past and prepare for the future, they implicate various relations including causal ones, and they affect intentions and decisions. They modulate emotions such as regret and relief, and they support moral judgments such as blame”. Empirical research has found that there can be different kinds of counterfactual thinking. For instance what is now called “episodic counterfactual thinking” refers to personal memories: “imagining alternative ways in which past personal events could have occurred but did not” and is allegedly different from semantic counterfactual thinking (De Brigard & Parikh, 2018). Moreover, it seems that counterfactual thinking has typicality effects (Catellani et al, 2008), and changes depending on whether it starts from a negative outcome (McCrea, 2007). Counterfactual thinking is affected by emotion (Atkinson et al, 2009). For instance, “repetitive counterfactual creation is one of the major forms of pathological rumination... resulting in extreme feelings of regret and disappointment.” (De Brigard et al, 2018, p.6). Besides, depressed people have functional differences in counterfactual thinking (Quelhas et al 2008). And it affects people with chronic pain (Andersson et al, 2006). Accordingly, De Brigard et al (2018, p.5) say: “These and related results suggest, therefore, that how plausible we think a counterfactual event is may depend on the extent to which we see it as immoral, abnormal, salient, or easy to imagine or whether it has been repeatedly simulated.”

(2b) If Oswald hadn't shot Kennedy, someone else would have.

Here is another:

(3) Shakespeare

(3a) If Shakespeare didn't write *Romeo and Juliet*, someone else did.

(3b) If Shakespeare hadn't written *Romeo and Juliet*, someone else would have.

From a semantic point of view (2a) and (3a) are pretty straightforward, while (2b) and (3b) are, if not anomalous, at least, strange. In order to analyze that "oddness" several proposals have been made, mainly using possible-worlds semantics, and therefore, pursuing truth conditions for those conditionals. For instance, Lewis's framework has been proposed to account for the idea that the closest world in which the indicative sentence (2a) is interpreted is such that a *common ground* is taken in which Kennedy was, actually, shot (cf. e.g. Williams, 2008). Truth conditions are important, but I am not going to focus on them here. My issue in this section is this: from an informational point of view, what makes the consequents of (2b) and (3b) strange?

To begin with, let's take another (apparently) unproblematic example:

(4) Tibbets

(4a) If Paul Tibbets didn't drop the atomic bomb on Hiroshima, someone else did.

(4b) If Paul Tibbets hadn't dropped the atomic bomb on Hiroshima, someone else would have.

There is no strangeness here. (4a) and (4b) are both acceptable. (2a), (3a), and (4a) seem analogous. In the common ground we take for granted that the antecedents are false, because in each case, what is denied is part of the common knowledge and official history. What are, then, the differences with (2b) and (3b)? Notice that in Tibbets's case, he was a pilot and part of a military corps, presumably as prepared for such a mission as any other pilot of that corps. Shakespeare and Oswald were not in analogous positions. Common knowledge says they acted as a single author and a single shooter. I will call the *tail* of information the class in which a subject is being included (predicated). In this sense, (4b) has a wide tail, because Tibbets is part of a military corps with other members. Now, what if the tail of information widens for (2b) and (3b)? (*wt* stands for 'widened tail'):

(2wt) Oswald is a member of a conspiracy. As such, he is one of many possible shooters.

(3wt) Shakespeare is really part of a group of authors and ghostwriters operating under the name "Shakespeare", with different contributors (Lord Earl, Francis Bacon, among others), [*à la* Nicolas Bourbaki].

If we read again (2b) and (3b) with (2wt) and (3wt) in mind, they become, if not "a matter of course", at least, plausible. One way to contrast the proposal is to 'contract' an already wide tail. In this case, we need to contract the tail of (4) (*ct* stands for 'contracted tail'):

(4ct) All Tibbets' companions are rebellious, non-compliant, pacifist military men.

If we read again (4a) and (4b) in the framework of (4ct), their consequents, if not impossible, at least lose the plausibility they have with their usual tails.

However, there is an asymmetry. Widening the tails of (2) and (3) affects the plausibility of both (2b) and (3b) and (2a) and (3a); contracting the tail of (4), however, affects (4b) and (4a) differently. Let's see this more closely.

In the Oswald case the widening is compatible with the shooting of Kennedy both in (2a) and (2b). In (2a) because where there is a shot (typically) there is a shooter (it is very implausible that, for instance, lightning struck a gun store, causing an accidental shot to hit Kennedy's vehicle). Thus, from the information that Kennedy was shot we can 'safely' say that there was a shooter ('safely' means here: as a robust default). And that shooter might or might not have been part of Oswald's conspiracy (maybe we can imagine, besides, various conspiracy groups). But in the case of (2b) the common ground information about the shooting of Kennedy is affected in another way. First, the information about Oswald, unlike (2a), does not deny his actions, it 'suspends' them: his shooting-action is bracketed. But, being a member of a conspiracy against Kennedy, the widened information is compatible with the idea that one of his companions would have shot, with the same result for Kennedy.

The Shakespeare case is analogous: the widening does not affect the common ground information about the content of *Romeo and Juliet*, and it is compatible with (3a) and (3b). But in (3a) Shakespeare's quill is inert, while in (3b) it may be available to be used.

Now, in the Tibbets case, although shortening the information does not affect our common ground information about the dropping of the atomic bomb on Hiroshima, it makes the compatibility between antecedent and consequent implausible in (4b) (as expected, if it is the reverse of widening). Effectively, what I said about the "shooter"- "shot" relation applies here as well: if there is a "drop", we can say 'safely' that there is a "dropper". So, even if the contracted information says that only Tibbets could have dropped the bomb, in (4a) we accept that the bomb was dropped, and then we accept the consequent: there must be another dropper (and probably none of Tibbets' companions). On the other hand, in (4b) if the contracted information says that only Tibbets could have dropped the bomb, then the consequent is unacceptable because nobody else would have done it¹³.

Now, if we see (1a) and (1b) again, we can check that the tail of (1) is wide enough already. But without additional information, we can only say that (1) is more vague than (4).

For the sake of argument, let's say that the widening-contraction of information tails has these effects on common-ground knowledge. We should ask why, anyway, they are not homogeneous. The widening in Oswald's case seems more plausible than in the Shakespeare case. The examples given treat composing a work of art and shooting a gun in a similar way. However, to create a masterpiece seems more 'unique' than to shoot a

¹³ Information is susceptible of being re-widened or re-contracted. For instance, we can imagine that Oswald is part of a conspiracy, but that all his companions are not willing to shoot, because of cowardice.

gun. Common knowledge says that becoming an accomplished writer requires a lot of talent, work, time, etc.; and even more for the most outstanding of the accomplished writers, so much so that for many such a talent is a real mystery (the Myth of Genius). On the other hand, training and resources for becoming a shooter, even an above-average one, are not so demanding.

It seems to me that this phenomenon can be related to some features of *generics*. Here I return to what Woods said about it. In Gabbay & Woods (2003, p.17) we can see the idea that a generalizer typically projects from samples to a generic proposition and not to a universally quantified conditional proposition. In the latter case, one counter-instance is enough for its refutation; but the first case can hold even in the face of multiple counter-instances. In this sense, a universally quantified proposition such as “all men are liars” is *brittle* because if we find just one man who is not a liar, it will be ‘ruptured’: it forces us to abandon it. On the other hand, there are *elastic* propositions, such as “birds fly” which can ‘resist’ counter-examples. For instance, from “Tweety is a bird” jointly with “birds fly” we can conclude “Tweety flies”. But if we find out that Tweety does not fly (the reasons can be many: it was born without wings or it is a penguin), that does not urge us to abandon the proposition “birds fly”. The inference by which a generic proposition is obtained is called “generic inference” (Gabbay & Woods, 2005, p.24). In (2013, p.244), Woods reflects once again about generics and proposes a sort of classification (my words, not his). In this classification about generics, Woods proposes three different levels, without a sharp demarcation between them:

Generic Generalizations, which have high nomological force, and whose counter-instances are ‘anomalies’. Think, for instance, of Ivan, the three-legged tiger. We can say that given the generic “tigers are four-legged”, there is something wrong with Ivan.

Adverbially Qualified Generalizations (non-quantifications –as in “typically”, “normally”, or “characteristically”), which have intermediate nomological force, and whose counter-instances are ‘oddities’. Think, for instance, of Guido, whose father is worried because of his non-Italian behavior (he is not a festive person, he does not like wine or pasta). It is not that there is something wrong with Guido, but certainly he does not go with the flow.

Non-Universal Plural Quantifications, which have low nomological force, and whose counter-instances are ‘rarities’. Think, for instance, of Tweety the non-flying bird. Although there are at least 40 species of non-flying birds, most birds fly. Thus, certainly, there is nothing wrong with Tweety, but it is not very common.

These three ‘levels’ of the nomological force of generics (and resistance to negative instances) seem to me to correlate with four features: (1) typicality effects (Lakoff, 1987), (2) level of entrenchment (Langacker, 2008), (3) admissibility (Niño, 2015, see above), and (4) level of vagueness-specificity (Short, 2007).

Typicality

First, it seems to me that the discussion about generics fits well with what cognitive linguistics literature calls *typicality effects* (cf. Lakoff, 1987) for the use of categories, particularly, stereotypes. Briefly, when we think of a bird, we think first of a feathered

flying animal, like a duck, or an eagle, and less readily of a penguin or an ostrich. Or in thinking about a means of transportation, we think first of a car, bus, or train, and less readily of an elevator or a skateboard, even if *stricto sensu*, an elevator is a means of transportation. So, just as generics belong to an axis that goes from the brittle to the elastic, so the typicality effects go from the typical to the strict. Studies have shown that categories for actions, colors, animals, among others have typicality effects (cf. Evans & Green, 2006; Ungerer & Schmid, 2006).

Typicality effects consist in the cognitive effects that the use of concepts has, given that we think and reason more easily by thinking first of a prototypical exemplar rather than one that is less typical: in other words, it is cheaper for our cognitive economies. In this sense, typicality effects are related to what Ronald Langacker calls *entrenchment* (1987, p.59-60, 380; 2008, p.16-17, p.220), that is to say, the more routine the use of a concept is, the more entrenched that concept becomes and the more easily it is activated. The cognitive use of stereotypes seems to be congenial with the cognitive economies required for practical agents. Now, if we accept that not only categories but also common knowledge has “typicality effects”, then, I would like to add, tail-widening or tail-contracting also have them; and that would contribute to an explanation of the range of plausibility of the widening-contraction operations¹⁴.

Admissibility

Orthogonal to the nomological force of generics is the *admissibility spectrum*, mentioned already in section 2. It consists in the level of admissibility an agent can have concerning a certain content, in the Whewellian sense of being consilient with other contents and the degree of ‘fitness’ it has in relation to other contents of his/her conceptual structure. In the case of belief¹⁵, admissibility becomes *credibility*, and it covers two axes: On the one hand, *verisimilitude*, if the ‘factual’ content fits in with other beliefs (in a very broad sense of ‘factual’: it is also a ‘fact’ in fiction that Holmes and Watson are friends, but it is not that Watson constructs a proton bomb and detonates it; and it is a ‘fact’ in mathematics that there are infinities greater than others. An agent can have different universes of discourse in his/her mind, with different ontologies). On the other, *acceptability*, if the axiological content (moral, deontic, judicial, political, veritative) can be –or even, should be– (more or less) approved or accepted by the agent. Sometimes credibility can be deliberately relaxed, in order to obtain new insights (as in fictional literature) or for the sake of fun: we admit the Shrek world where different fairy tales converge in one universe of discourse just for the entertainment it provokes.

Thus, any new content can be ‘evaluated’ in a range of admissibility, but older contents can be ‘re-evaluated’ also. Maybe at some point some are no longer admissible, maybe we consider them now as errors (‘factual’ mistakes), regrets (‘unacceptable’ mistakes), etc. In this sense, we have a ‘no-longer-admissibility’ system, which is the counterpart of

¹⁴ Notice, by the way, that if this is true, fans of conspiracy theories do not flinch at common knowledge implausibilities; and just as academic skeptics, they are self-motivated to remain so.

¹⁵ In the case of desire, for instance, the scope of this discussion is different.

admissibility. In this sense, in our web of beliefs, there are certain contents that range from ‘sturdy’ to ‘weak’, because they present more or less resistance to their ‘movability’. For instance, at this moment I am convinced that my car is in the garage. With that expectation I go there and... I don’t see it. Between “Maybe I’m wrong, maybe I entered the wrong garage” (as when Davidson entered the wrong home) and “I was robbed!” there are many intermediate possible states. What I want to point out is that the first response shows a certain sturdiness and the second a certain weakness. Neither of them, by itself, guarantees truth. Maybe both of them are wrong: the car is there but my visual system fails. Maybe I was wrong from the beginning: my car was never in the garage, or even, I don’t have a car. The difficulty/ease in removing a content correlates with the difficulty/ease in admitting a previous one. This gives us another lesson: proportional to our expectation is the level of surprise. Until 1697 people thought all swans were white. But Englishmen discovered that year, in Australia, black swans (*Cygnus atratus*). So, for native Australians there is no surprise in seeing a black swan. For other people the surprise is enormous: “It cannot be! It must be a fake!”. And now, if Jane says to Mary: “A new swan arrived at the zoo and is white”, Mary can think that the qualification “and is white” is not trivial, because it may well have been black, even if her own initial expectation was, before the qualification, that –indeed– it was white¹⁶. And so, in abduction, “surprise” moves along an axis from ‘no-surprise-at-all’ to ‘this-can’t-be’. As a counterpart, then, “matter-of-course” also moves from ‘obviously!’ to ‘this-is-unintelligible!’

Notice that the admissibility condition has some bearing on the structure of categories (prototypical and radial). Typically an “aunt” has ties with her nieces and nephews, full of affection, care, and love. But not Harry Potter’s aunt. However, this does not rupture our generic for “aunt”: her conduct toward Harry is verisimilar, and so, it also tells us that she is a bad (unacceptable) aunt.

So, the way contents are formed and processed (and in this, both entrenchment and admissibility are crucial), has a bearing on the rupturability of belief, because rupturability depends on a bottleneck in the information structure; and this structure is constrained by entrenchment, admissibility, and level of vagueness, which is our next topic.

Vagueness & Specificity

This point is highlighted by Thomas Short in his discussion of Peirce’s hypostatic abstraction (cf. Short, 2007, Ch. 10). Here we understand *vagueness* as a lack of specificity and not as fuzziness; it implies lack of information about *that to which* the vague concept is being applied. For instance, we can say that “madness” is *that which* makes a person’s behavior strange and his/her mental health troubled. And about this point, say, psychiatrists agree. In this sense, “madman” is equally vague for them and for the layman. But psychiatrists do not agree on how to *specify what* causes the disorder. Is it, say, because of unconscious repression or a problem with chloride channels in the membrane of certain neurons, or something else? But even if some agree that the problem is in the

¹⁶ Maybe I’m not saying anything new: belief change is more than expansion, retraction, and revision.

chloride channels, it is possible to *specify* further the hypothesis, to establish if the chromosomes responsible for this are *those which* express this or that protein.

A crucial point concerning the degree of vagueness or specificity is that a certain level of vagueness allows the generation of a *common framework*, which in turn, has an impact on the generation of a *common ground* (see above). Think, for instance, of the context in which the dispute emerges about the explanation of heat (cf. Short, 2007, p.270-271). On the one hand, there were those who argued that *that which* explains the heat phenomena is a substance, the 'caloric'. On the other hand, there were those who argued that *that which* explains the changes of temperature was something like the 'movement of particles'. But both of them shared the idea that there was something, something that could explain phenomena such as the change of temperature, the transfer of heat between two bodies with different temperatures, the boiling of water, the production of combustion, etc. That is, the 'calorists' and the 'kinetic theorists' agreed that *there was something* that produced these phenomena. Where they differed was in the way of understanding what that something consisted of: matter or movement¹⁷.

Finally, I do not want the vagueness/specificity of information to be confused with its precision/imprecision. Information is more or less *precise* depending on the agenda that information is supposed to help to close. (That's why in the first section I spoke of "standards of precision"). In that sense, 'precise' or 'imprecise' are evaluations of the role certain information plays, while vague/specific is a feature of information, depending on the structure it has. For instance, the information that at this moment "Bruce Kent is in Washington" will be judged as precise enough (or not), depending on whether the interest in that information is related to the purpose of his girlfriend to know if her boyfriend has already arrived in that city, or if the interest comes from the FBI, because its officials suspect that Bruce Kent carries a nuclear weapon and want to arrest him. In the first case, it is possible that the information of the assertion is considered sufficiently precise; while in the second it surely is not. Very imprecise information risks being insufficient to close the agenda-in-course. Very precise information risks being too demanding to be learned, retrieved or processed, and so becoming cognitive noise. Therefore, precision has some bearing on relevance, which is our next theme.

Relevance, Centrality and Salience

According to Woods, *relevance*:

is defined for quadruplets $\langle X, I, B, A \rangle$ in which X is a practical agent, I is some information, B his background information, and A is his cognitive agenda. Thus information is relevant for an agent with respect to a given agenda when, relative to his background information the incoming information is processed by the agent in ways that produces a state of mind that advances or closes that agenda... In its most basic sense information is relevant when it is helpful. A further feature of agenda relevance approach is that relevance is a causal relation, rather than a semantic or probabilistic one. Information is relevant for a cognitive agent when in processing it his mind is caused to change in ways that are cognitively helpful to him. (Woods, 2013, p.28-29).

¹⁷ In this sense, the Kuhnian incommensurability of theories (Kuhn, 1970) seems to lack force.

This means that relevance is important for the closure of agendas-in-course and planned agendas, and being a causal relation is compatible with the CRM of knowing. I think Ronald Langacker's *Cognitive Grammar* has designed some analytical tools that can help us examine the cognitive phenomena activated in this process, if relevance is causal. In his cognitive approach to grammar, Langacker has proposed a very interesting theoretical framework in order to understand the structure and use of our conceptual architecture. According to Langacker (2008), language use has the structure of attention (figure/ground); and using a *domain* (i.e. a set of contents available to an agent in his/her memory) mirrors that structure, which he calls *profile/base* (if the concept refers to an object), and *trajector/landmark* (if the concept refers to a relation). For instance, in order to understand "hypotenuse" (see Figure 1, Langacker, 2002, p.3-4), it is necessary to understand "right triangle". Here "hypotenuse" is the profile and "right triangle" is the base.

Figure 1.



Profile/Base is a relative structure. In using "elbow", "elbow" is profiled against the base "arm"; but in using "arm", "arm" is profiled against the base "human body". Most contents are composed of a matrix of domains. For instance, "knife" could refer to various domains: space, form, cutting, the bodily actions associated with different uses, its inclusion in a cutlery setting, the place where we can buy it, etc. All those specifications are different in their *centrality*, that is, the likelihood that a domain becomes active on a given occasion of the expression's use (Langacker, 2002, p.4). Now, "being used for a purpose presupposes purposeful action" (Short, 2007, p.111). Thus, use presupposes the purpose of that use. In this sense, an expression's use presupposes what an agent is trying to achieve. If we understand this from an agendive perspective, centrality consists in the degree of contribution the content of an expression has in relation to an agenda-in-course. In other words, its contribution to the agent's degree of precision. Of course, different contents may have a similar centrality (think here of synonymy). In this sense, centrality is also related to the scope conditions of an agenda-in-course: An agent may have at his/her disposal different contents with similar centrality, if he/she is competent enough so that, on the one hand, he/she has the information; but on the other, he/she can recruit and use that information at the appropriate time. I am aware that relevance is a causal relation in Woods (2013, p.29) and centrality is a semantic feature in Langacker. However, what I am proposing is that the presence and processing of 'central' information contributes to the advancement of a cognitive agenda, and if helpful (given the conditions of a CRM), becomes relevant.

A second notion of Langacker's Cognitive Grammar I want to mention is *Saliency*. (Cognitive) Saliency consists in a certain portion of a domain being activated (active zone) and loaded into working memory, thus becoming part of the center of attention. Since the

concepts that are already active require fewer cognitive resources, a greater degree of cognitive salience is correlated to greater ease of activation facilitating a greater cognitive economy. By contrast, inactive concepts will be less salient (Schmid, 2007: 119). Three comments are in order here. First, salience processes probably occur 'down below', although their effects appear in the center of attention. Second, ease of activation is related to the entrenchment of a content: the more entrenched a content is, the faster and more automatic its activation will be. Insofar as entrenchment is an effect of habituation (due to repeated use), it also is related to the level of expertise an agent has (fluency) (cf. Langacker, 2008, p.16-17, p.220). Third, insofar different salience levels affect what can appear in the attentional focus, salience can affect the drawability of consequences: you will not infer a conclusion from a premiss that you have not activated (mistakes of omission, cf. Woods, 2013, p.26). One additional word: salience and centrality are very different: a diversion is something that is present in your attention (salience), but contributes very little to your interests (centrality).

5. Conclusion

The project of a naturalization of logic is not only interesting, but necessary in order to give an accurate account of real reasoning performed by real agents in real time. It should be empirically informed, particularly by the discoveries of cognitive science.

However, in order to achieve its goal, it is necessary to 'naturalize' other things. I think the most necessary starting point is the naturalization of meaning. Some advances have been made so far, but it seems we are witnessing just the birth of it. Naturalization of meaning involves the naturalization of meaning generation (imagination and creativity). This has some bearing, for instance, on the not yet developed relation between proposals such as conceptual blending (Fauconnier & Turner, 2002; Turner, 2014) and the generation of the second premiss of abduction. And this could have some bearing on a practical logic: Maybe there is a sense in which, just as Peirce says we cannot doubt anything at will, because there are things we do not even imagine can be doubted, we do not draw consequences just because we do not even imagine they can be drawn. So, lack of imagination can become an error of reasoning.

Suppose the above has some bearing. How far, then, is it necessary to take psychologization in logic for the naturalization project? We do not know yet.

Can a naturalized logic be neutral concerning the dispute between computationalists and embodimentalist in philosophy of mind?

References

Adams, E. W.

1975. *The Logic of Conditionals*. Dordrecht: Reidel.

Atkinson, Lisa; Bell, David & Feeney, Aidan

2009 "The relationship between counterfactual thinking and emotional reactions to event outcomes: Does one account fit all?". *Psychonomic Bulletin & Review*, **16** (4), 724-728. doi:10.3758/PBR.16.4.724

Bennett, Jonathan

2003. *A Philosophical Guide to Conditionals*. Oxford: Clarendon Press.

Byrne, Ruth

2016 "Counterfactual Thought". *Annual Review of Psychology*. **67**:7.1–7.23. doi: 10.1146/annurev-psych-122414-033249

Catellani, Patrizia; Alberici, Augusta I. & Milesi, Patrizia

2004 "Counterfactual thinking and stereotypes: The nonconformity effect". *European Journal of Social Psychology*. **34**, 421–436. DOI: 10.1002/ejsp.206

De Brigard, Felipe & Parikh, Natasha

2018. "Episodic counterfactual thinking". *Current Directions in Psychological Science*, 1-8. DOI: <https://doi.org/10.1177/09637214188065>

De Brigard, Felipe; Hanna, Eleanor; St Jacques, Peggy L. & Schacter, Daniel L.

2018. "How thinking about what could have been affects how we feel about what was", *Cognition and Emotion*, DOI: 10.1080/02699931.2018.1478280

Evans, Vyvyan & Green, Melanie

2006. *Cognitive Linguistics. An Introduction*. Edinburgh, Edinburgh University Press.

Fauconnier, Gilles & Turner, Mark

2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Gabbay, Dov & Woods, John

2003. *Agenda Relevance. A Study in Formal Pragmatics (A Practical Logic of Cognitive Systems, Volume 1)*. Amsterdam: Elsevier.

2005. *The Reach of Abduction. Insight and Trial (A Practical Logic of Cognitive Systems, Volume 2)*. Amsterdam: Elsevier.

Kuhn, Thomas

1996 [1970, 1962]. *The Structure of Scientific Revolutions*. 3rd ed. Chicago: The University of Chicago Press.

Lakoff, George

1987. *Women, Fire, and Other Dangerous Things*. Chicago: The University of Chicago Press.

Langacker, Ronald

1987. *Foundations of Cognitive Grammar*. Vol. 1. Stanford: Stanford University Press.

2002 [1991, 1st ed.]. *Concept, Image, and Symbol. The Cognitive Basis of Grammar*. 2nd ed. Berlin: Mouton de Gruyter.

2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.

Magnani, Lorenzo

2018. "The Urgent Need of a Naturalized Logic". *Philosophies*, **3**, 44, 1-16;
doi:10.3390/philosophies3040044

McCrea, Sean M.

2007 "Counterfactual thinking following negative outcomes: Evidence for group and self-protective biases". *European Journal of Social Psychology*. **37**, 1256–1271. DOI: 10.1002/ejsp.462

Niño, Douglas

2008. *Abducting abduction. Avatares de la comprensión de la abducción de Charles S. Peirce* (PhD Thesis, Unpublished). Bogotá: Universidad Nacional de Colombia.

2013. "Agentes y circunstancias. Consideraciones en torno a factores epistémicos y fiduciarios en la argumentación intercultural" (pp. 187–246). In Camilo Martínez (Ed.), *Cuatro aproximaciones al diálogo argumentativo intercultural*. Bogotá: Universidad Jorge Tadeo Lozano.

2015. *Elementos de semiótica agentiva*. Bogotá: Universidad Jorge Tadeo Lozano.

Niño, Douglas & Marrero, Danny

2015. "The Agentive Approach to Argumentation: A Proposal" (p.53-67). In: Frans H. van Eemeren & Bart Garssen (Eds.), *Reflections on Theoretical Issues in Argumentation Theory*. Dordrecht: Springer.

Peirce, Charles

1903. "The logic of abduction" (pp. 226–241). In The Peirce Edition Project (Eds.) *The Essential Peirce* (Vol. 2). Bloomington: Indiana University Press.

CP. 1994. *The Collected Papers of Charles S. Peirce*. 8 vols. Vols. 1-6 Edited by Charles Hartshorne and Paul Weiss; vols. 7-8, Edited by Arthur Burks. Cambridge: Harvard University Press. Electronic version by IntelLex Corporation, 1994. Followed by volume and paragraph.

MS. 1967-1971. *The Charles S. Peirce Papers*, 32 rolls of microfilms from the original, deposited at the Houghton Library, Photographic Service, Harvard University Library, Cambridge, Massachusetts. Numbering corresponds to the Annotated Catalogue of the Papers of Charles S. Peirce. Richard Robin. Amherst: University of Massachusetts Press, 1967 and/or "The Peirce Papers: A Supplementary Catalogue", Transactions of the Charles

S. Peirce Society, 7 (1971): 37-57. "ISP" refers to the manuscripts' page numbering by the Institute for Studies in Pragmaticism.

RLT. 1992. *Reasoning and the Logic of Things. The Cambridge Conferences Lectures of 1898*. Edited by Kenneth Laine Ketner and Hilary Putnam. London: Harvard University Press.

Quelhas, Ana Cristina; Power, Mick J.; Juhos, Csongor and Senos, Jorge

2008 "Counterfactual Thinking and Functional Differences in Depression". *Clinical Psychology and Psychotherapy*. 15, 352–365.

Quine, Willard V.O.

1969. "Epistemology Naturalized" (p.69-90). In: *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Schmid, Hans-Jörg

2007. "Entrenchment, Saliency, and Basic Levels" (pp. 117-138). In: *The Oxford Handbook of Cognitive Linguistics*. Dirk Geeraerts & Hubert Cuyckens (Eds.). Oxford: Oxford University Press.

Ungerer, Friedrich & Schmid, Hans-Jörg

2006. *An Introduction to Cognitive Linguistics*. 2nd ed. London: Pearson Longman.

Williams, J. Robert G.

2008 "Conversation and conditionals". *Philosophical Studies*, **138**: 211–223.

Woods, John

n.d. "Logic Naturalized" (preprint). www.johnwoods.ca

2007. "Eight Theses Reflecting on Stephen Toulmin." (preprint). www.johnwoods.ca

2007. "The Concept of Fallacy is Empty. A Resource-Bound Approach to Error" (p.69-). In: Lorenzo Magnani and Ping Li (Eds.) *Model-Based Reasoning in Science, Technology, and Medicine*. Berlin: Springer.

2011. "Recent developments in the logic of abduction". *Studies in History and Philosophy of Science* **42**, 240–244.

2013. *Errors of Reasoning. Naturalizing the Logic of Inference*. London: College Publications.

2017. "Reorienting the logic of abduction" (p.137-150). In: Lorenzo Magnani, Tommaso Bertolotti (Eds.), *Springer Handbook of Model-Based Science*. Dordrecht: Springer.

2018. *Truth in Fiction. Rethinking its Logic*. Cham: Springer.