The benefits of prototypes: The case of medical concepts

M. Cristina Amoretti*, Marcello Frixione*, Antonio Lieto°

* University of Genoa

° University of Turin and ICAR-CNR

pre-print version. To appear in Reti, Saperi e Linguaggi, The Italian Journal of Cognitive Sciences, 2017.

Abstract

In the present paper, we shall discuss the notion of prototype and show its benefits. First, we shall argue that the prototypes of common-sense concepts are necessary for making prompt and reliable categorisations and inferences. However, the features constituting the prototype of a particular concept are neither necessary nor sufficient conditions for determining category membership; in this sense, the prototype might lead to conclusions regarded as wrong from a theoretical perspective. That being said, the prototype remains essential to handling most ordinary situations and helps us to perform important cognitive tasks. To exemplify this point, we shall focus on disease concepts. Our analysis concludes that the prototypical conception of disease is needed to make important inferences from a practical and clinical point of view. Moreover, it can still be compatible with a classical definition of disease, given in terms of necessary and sufficient conditions. In the first section, we shall compare the notion of stereotype, as it has been introduced in philosophy of language by Hilary Putnam, with the notion of prototype, as it has been developed in the cognitive sciences. In the second section, we shall discuss the general role of prototypical information in cognition and stress its centrality. In the third section, we shall apply our previous discussion to the specific case of medical concepts, before briefly summarising our conclusions in section four.

1. Stereotypes and prototypes

In philosophy of language, the term "stereotype" was introduced by Hilary Putnam (1975). He famously used the notion of stereotype to account for a speaker's semantic competence, which is identified with her knowledge of the stereotypes associated with words. To put it briefly, a stereotype is "a standardized description of features of the kind that are typical, or 'normal', or at any rate stereotypical. The central features of the stereotype generally are *criteria* – features which in normal situations constitute ways of recognizing if a thing belongs to the kind" (1975, p. 230), that is, ways of categorising a particular thing. A stereotype is "a conventional […] idea […] of what an *X* looks like or acts like or is." (1975, p. 249). According to Putnam, a stereotype is thus an oversimplified theory of a word's extension (for example, being yellow and shining, for gold; being black-and-yellow striped, and having long teeth, for tigers) that all competent speakers should know, as it is the required minimum level of competence to communicate with others in a given linguistic community. However, knowing the stereotype does not suffice to know the meaning of a word or to fully grasp the concept associated with it, as they do not determine reference in the appropriate way: pyrite is yellow and shining without being gold, and some types of gold are not yellow; wax-made tigers are not tigers, while albino tigers are. Yet, some members of the linguistic community, that is the experts, are more competent than others at determining the features that a thing should have in order to really belong to a particular kind and to be properly named with a certain word: for example, having 79 as atomic number for gold; being a member of *Panthera tigris* because of specific biological features. Anyway, the stereotype associated with a word describes the typical thing associated to that word, or what the community takes it to be: the stereotype of gold describes the typical instance of gold; the stereotype of tiger describes the typical tiger; and so on. On this perspective, stereotypes are important tools that help us crucially in our everyday processes of categorisation; however, they merely work as heuristics that do not always allow for an accurate categorisation, as they are not sufficient, by themselves, to pick out the right extension of a term.

The notion of stereotype is often compared or confused with that of prototype. This latter notion has been introduced by Eleonor Rosch to account for the typicality effects that are exhibited by many common-sense concepts, such as fruit or bird.

Eleanor Rosch's psychological experiments (Rosch 1975; Rosch and Mervis 1975) showed how common-sense concepts do not obey the requirements of the so-called classical theory, according to which concepts must be defined in terms of sets of individually necessary and jointly sufficient conditions. The majority of common-sense concepts cannot be defined in terms of a classical definition (and even if for some concepts such a definition is available, subjects do not use it in many cognitive tasks). Rather, concepts exhibit *typicality effects*: some members of a category are considered better instances than others. For example, a robin is considered a better example of the category of birds than, say, a penguin or an ostrich. More central instances share certain typical features (e.g. the ability of flying for birds, having fur for mammals) that, in general, are neither necessary nor sufficient conditions. Rosch explains such effects by hypothesising that concepts are mentally represented in terms of *prototypes*, where a prototype is a mental representation of a typical member of a given category. A prototype is an ideal entity that does not necessary coincides with any real individual.

Obviously, the notion of stereotype and that of prototype have many relevant features in common and are, thus, very similar. However, they also differ in some important respects. On the one hand, the notion of stereotype was developed by Putnam in close relationship with his particular conception of semantic meaning. The notion of stereotype thus belongs to philosophical semantics. On the other hand, the notion of prototype was developed by Rosch to account for typicality effects. In this respect, the notion of prototype belongs to empirical psychology and is not committed to the specific semantic theory proposed by Putnam. In order to avoid possible confusions, in what follows, we shall only use the term "prototype", as the arguments we discuss mostly come from the cognitive sciences.

2. The benefits of prototypes

Typicality effects in categorisation and, in general, in category representation are crucial for human cognition. Under what conditions should we say that somebody *knows* the concept DOG (or, in other terms, that he or she possesses an adequate mental representation of it)? It is not easy to say. However, if a person does not know that, for example, dogs usually bark, that they typically have four legs, that their body is covered with fur, that, in most cases, they have a tail and that they wag it when they are happy, then we should probably conclude that this person does not grasp the concept DOG. Nevertheless, all of these pieces of information are neither necessary nor sufficient conditions for being a dog. In fact, they are traits that characterise dogs in typical (or prototypical) cases.

The concept DOG is not exceptional from this point of view. The majority of everyday concepts behave in this way. For most concepts, a classical definition in terms of necessary and sufficient conditions is not available (or, even if it is available, it is unknown to the agent). Moreover, it may be that we know the classical definition of a concept, but typical/prototypical knowledge still plays a central role in many cognitive tasks. Consider the following example: nowadays most people know the necessary and sufficient conditions for being WATER: water is the chemical substance whose formula is H_2O, i.e., the substance whose molecules are formed by one atom of oxygen and two atoms of hydrogen[1]. However, in most cases of everyday life, when we categorise something as WATER, we do not take advantage of this piece of knowledge. Instead, we use prototypical traits, such as the fact that (liquid) water is usually a colourless, odourless and tasteless fluid.

Human beings have a strong inclination to reason in prototypical terms. Consider, for example, the so-called Linda problem and the conjunction fallacy (Tversky & Kahneman 1983). In a well-known experiment from the psychology of probabilistic reasoning, subjects are given a description of a person named Linda that perfectly fits the stereotype of a feminist activist ("Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations", Tversky & Kahneman 1983, p. 297). Then, they are asked to judge whether it is more likely that Linda is (a) a bank teller or (b) a bank teller *and* a feminist. The majority of subjects choose (b), without realising that being a feminist bank teller is in any case more demanding (and therefore less probable) than simply being a bank teller. Indeed, the conjunction fallacy consists in failing to realise that a conjunction is always less probable than its conjuncts. From the standpoint of a theory of reasoning, this is a paradigmatic case of a systematic error. However, if we consider the problem from the point of view of a theory of concepts, the conjunction fallacy can be interpreted as an example of the strong tendency of human subjects to resort to prototypical information in categorisation, even when this is not appropriate (at least from the point of view of those who devised the experiment): Linda is categorised as a feminist because she perfectly fits the prototypical traits of a feminist (while she does not in the least fit the prototypical traits of a bank teller). So, we could see the "error" as originating from a process of categorisation based on prototypical knowledge.

The use of prototypical knowledge in cognitive tasks such as categorisation is related to the constraints that concern every finite agent that has limited access to the knowledge relevant for a given task. In most cases, cognitive processes based on typical knowledge are fast, automatic and cognitively undemanding (and are presumably homogenous to the processes employed in similar tasks by non-human animals). Consider, for example, the following variant of the Linda problem. Let us suppose that a certain individual named Pippo is described as follows. He weighs about 200 kg, and he is approximately two meters tall. His body is covered with a thick, dark fur, he has a large mouth with robust teeth and paws with long claws. He roars and growls. Now, given this information, we have to evaluate the plausibility of the two following alternatives:

a') Pippo is a mammal;

---

[1] This, in strict terms, is not a *definition*, since it is a piece of a posteriori knowledge. But, for our present purposes, this is not relevant.

b’) Pippo is a mammal, and he is wild and dangerous.

Which is the “correct” answer? According to the dictates of the normative theory of probability, it is surely (a’). However, if you were to encounter Pippo in the wilderness, it would be more appropriate to infer (b’).

This is an example of a default inference, in which a cognitive agent “jumps to a conclusion” on the basis of incomplete information. Typicality effects in categorisation and default inferences are closely related. For example, birds, by default (or prototypically), are able to fly; therefore, given that Tweety is a bird, you can jump to the provisional conclusion that Tweety is able to fly. Of course, you must be prepared to withdraw this conclusion should Tweety turn out to be atypical in his flying abilities (because, for example, he is an ostrich, a penguin, or a featherless chick). Such an inference from incomplete premises is an example of what is called non-monotonic reasoning (see e.g. Koons 2013).

Default values that allow the drawing of non-monotonic inferences are akin to prototypical knowledge. In other words, associating prototypical information to concepts turns out to be useful in order to make defeasible inferences starting from incomplete knowledge.

Default information is involved in countless cognitive tasks, such as reasoning on actions and events. Furthermore, consider the default assumptions that need to be built into the visual system in order to solve the ill-posed problem of reconstructing the properties of distal stimuli starting from incomplete proximal information. As such, it is likely that default “inferences” are exploited at many different levels of the cognitive architecture in completely independent components.

The existence of typicality effects is probably the most solid result achieved in the psychological research on concepts. Typicality effects in categorisation and, in general, in category representation are crucial for human cognition. However, it can be argued that typicality effects, far from being a symptom of some homogeneous cognitive structure, are more plausibly the consequence of some “ecological constraints” acting on the mind (Frixione 2013). In other words, our way of categorising the world does not exhibit typicality effects because our mind is structured in a particular way; rather, typicality effects are determined by the fact that the world interacts in a certain way with our mind.

This could have important consequences for the status of the notion of “concept” within the sciences of the mind. Since “conceptual” abilities such as categorisation are crucial for many cognitive tasks, it is plausible that different components of the cognitive architecture developed similar behaviours with respect to different tasks and different domains. Since these components are subject to similar constraints (such as limited access to information), it should not be surprising that this process should converge on similar solutions. Typicality effects in categorisation could be an example of such a “convergent evolution” within the mind.

Moreover, default reasoning is a widespread cognitive phenomenon that is likely to be determined by ecological constraints, namely by the need for cognitive agents to face various inferential problems with limited access to relevant knowledge.

In other words, it can be hypothesised that the emergence of typicality effects is in some sense analogous to convergent evolution in living beings: in some cases, species of different (and remote) lineages evolved similar characters as a response to similar ecological conditions. An example is the hydrodynamic shape evolved by pelagic fishes, dolphins and ichthyosauri. A further example is offered by the category of the so-called fat plants (or succulent plants), which constitute a heterogeneous group, including phylogenetically distant organisms, but which share convergent adaptations to similar ecosystems. For example, consider the spines and succulent tissues developed in such phylogenetically remote families as Cactaceae and Euphorbiaceae.

So, typicality effects, far from being a symptom of some homogeneous cognitive structure underlying categorisation, could, more plausibly, be the effect of some "ecological constraints" acting on the mind. In other words, our way of categorising the world does not exhibit typicality effects because our mind is structured in a particular way; rather, typicality effects are caused by the fact that the world interacts with our mind in a certain way. Typicality effects emerge as a result of the constraints that the environment imposes on any (finite) agent interacting with it. As such, they do not reveal any particular, specific feature of the human mind: all cognitive agents that have to survive in their environment would develop categories with a similar structure in order to solve similar problems. Obviously, this does not say very much about *how* these problems have been solved by the human cognitive system, nor does it involve some common cognitive mechanism underlying typicality effects. If this position is tenable, then the role of the concept of "concept" in cognitive science should be similar to the role of the concept of "fat plant" in botany: it should be of some utility in certain cases, but it needs not correspond to a genuine taxon (i.e., to a genuine kind).

In order to explain the "ecological" character of the constraints at the basis of prototypical behaviour, let us consider the following cases, which are in some sense paradigmatic. It may happen that, in order to accomplish their purposes, cognitive agents have the need to classify their environment according to categories that:

i. are intrinsically fuzzy (or vague); or

ii. do not correspond to an underlying common nature (e.g. to natural kinds); or

iii. correspond to substances whose essence is (totally or partially) unknown.

Point i. includes the case of categories that establish arbitrary discontinuities in phenomena that are intrinsically continuous. A paradigmatic example of this kind is colours: the chromatic spectrum is a continuum, and any way to categorise colours would give rise to vague concepts. Such categories cannot have clear-cut boundaries. Consequently, any chromatic category will exhibit better and worse exemplars, and forms of typicality effects will arise.

Point ii. includes cases in which typicality effects are determined by the fact that a category need not "capture an essence", i.e., it does not "carve the world at its joints". In order to satisfy their needs, cognitive agents are often interested in grouping together entities in ways that do not correspond to the "deep structure" of the world. Paradigmatic categories, in this sense, are categories that include natural entities, which, however, are grouped together according to criteria that, being determined by the agent's needs and purposes, are extrinsic to the organisation of the natural world. Good examples are food categories. Consider such lexical categories as VEGETABLE, FRUIT (as contrasted to VEGETABLE) or SEAFOOD. VEGETABLES, for example, include quite heterogeneous parts of vegetal organisms such as seeds, fruits, leaves, flowers, roots and branches. In the category of SEAFOOD almost all animal phyla are represented: mainly Chordata (Fishes), Mollusca and Artropoda (Crustacea above all), but also Echinodermata and Coelenterata[2].

It could be argued that, even if such categories do not correspond to natural kinds from a strictly taxonomical point of view, they could rather be considered "substances" from a biochemical standpoint: they include biological tissues that can be assimilated by the human organism. However, this does not constitute a sufficient condition. For example, in order to consider some stuff as

[2] Similar examples involving the fruit category have also been advanced by Borghini (2016).

VEGETABLE or SEAFOOD, further constraints must be satisfied. Its taste must not be disgusting, it must not be dangerous due to the presence of poisons or parasites, it must be easy enough to collect in satisfactory quantities, and so on (leaving apart cultural and anthropological constraints, maybe partially motivated by practical considerations, such as the rules of *kasherut* for Jews or the prohibition of cannibalism – cfr. Harris 1985). For example, *Amanita caesarea* is a rather atypical (not even being a plant) example of a vegetable, however, the poisonous *Amanita phalloides* cannot be considered as a vegetable, despite the biological similarity of these two species. Analogously, *Littorina neritoides* is not categorised as SEAFOOD, due to its small size, while the zoologically similar *Littorina littorea* counts as a good example of this category.

Case iii. includes those categories that correspond to an underlying nature that is (totally or partially) unknown to a cognitive agent. Consider the example of WATER mentioned above. In such cases, the agent adopts stereotypes (in the sense of Putnam 1975) that are partial and fallible, and which, therefore, give rise to uncertain and dubious cases that better fit some exemplars of a substance than others.

From a cognitive point of view, iii. is not dissimilar to the case of categories for which a classical, clear-cut definition is available, which, however, is unknown by the cognitive agent. Imagine an undergraduate student who knows that some members of the faculty are full professors, without exactly knowing how the concept FULL PROFESSOR is defined. This student would probably develop a “stereotype” that will help him in dealing with the academic world: full professors are usually older than their colleagues, they have greater authority and power, sometimes they are more erudite and have greater scientific prestige, frequently they are more arrogant, and so on.

In cases ii. and iii., typicality effects do not depend (only) on the vagueness of the employed criteria, but also on the “bad correspondence” between identification criteria and the world. Criteria are intrinsically imprecise, and give rise to dubious cases. In any case, all the above-mentioned cases give rise to typicality effects.

A further aspect deserves some consideration. It is well known from psychological experiments that human behaviour also exhibits typicality effects in the case of fully classical categories (examples can be found in arithmetic or geometry). This could be used against the thesis that typicality effects depend on ecological constraints and do not correspond to any specific feature of the cognitive representation of categories. The fact that people even exhibit prototypical behaviours for classical categories could be explained in terms of some deep feature of human conceptual representation. However, this is not necessarily the case. We can account for typicality effects for classical categories by adopting a distinction similar to the core/identification procedures distinction proposed by Miller and Johnson-Laird (1976) (on this point, see for example Armstrong *et al.* (1983) and Rey (1983)). There are cases in which the classical definition of some category (its core) is well known but where, in many practical cases, it cannot be applied and, instead, people must trust fallible criteria. For example, even if I know perfectly well what a *grandmother* is, when I categorise somebody as a grandmother I can seldom use the definition and must often rely on fallible criteria: grandmothers are old women with white hair, they are kind to children and so on. In some sense, this example is similar to case iii. above. It is also likely that, in the case of formal (mathematical) concepts, typicality effects probably depend on identification procedures. Let us consider even numbers. We tend to consider number 4 as a more “typical” specimen of this category than, say, 351,274,195,336. This is probably due to the fact that small even numbers (2, 4, 6, 8…) are presumably explicitly stored in our memory in some sort of lookup table, and are therefore easy to check and generate. For larger numbers, such as 351,274,195,336, we need a procedure ("check the last figure on the right") that, albeit simple, is more demanding (at least because it requires the procedure for small numbers as a part of the whole process). Compare this with the case of prime numbers: small prime numbers (2, 3, 5, 7, 11…) are explicitly stored in memory, and we can recognise them at a glance, whilst for larger numbers (e.g. 2,750,159) a rather complicated and time-consuming procedure is needed.

That typicality effects do not correspond to some specific feature of mental organisation is supported by the fact that, in some categorisation experiments, people dealing with abstract stimuli that have been carefully prepared by experimenters to favour the emergence of typicality effects, prefer to individuate classical categories (Medin *et al.* 1987). (In some sense, these results show that our mind has no particular bias against classical concepts and definitions. Rather, the problem is that classical concepts are not particularly useful for an agent with limited access to the world.)

The hypothesis that typicality effects could not correspond to some specific mental structure or to some specific feature of the cognitive system is corroborated by the observation that, in computational artificial systems, forms of "prototypical behaviour" can be obtained using very different computational devices. Typicality effects can be modelled by adopting, for example, some non-monotonic logical formalism (see Giordano *et. al* 2013); symbolic models of associative memory, such as frames (Minsky 1975) or inheritance networks (Horty *et al.* 1990); artificial (supervised or unsupervised) neural networks (see e.g. Socher *et al* 2013, Blouw *et al* 2016); or conceptual spaces (see e.g. Gärdenfors and Williams 2001, Lieto *et al.* 2015).

3. A case study: the concept of disease

Focusing on a contentious and debated concept in philosophy of medicine, that of DISEASE, we shall exemplify how prototypical concepts have a positive and important role from a practical point of view, in the domain of clinical medicine, as they may be very useful in everyday and clinical situations. At the same time, we shall maintain that their "negative" side, that is, the fact that prototypical concepts may result in incorrect categorisations, makes them compatible with the possibility that a classical and "conceptually-clean" definition of the concept of DISEASE, in terms of necessary (and sufficient) conditions, might eventually be found from the perspective of theoretical medicine.

To put it differently, we wish to show that, even if the prototypical DISEASE may eventually lead to categorisations considered wrong from the point of view of theoretical medicine, it is still essential for promptly and reliably handling most ordinary situations and, thus, guiding practical or clinical medicine. For this reason, the prototypical concept of DISEASE can be regarded as highly valuable and beneficial. However, as a prototypical concept does not pretend to exactly characterise the concept of DISEASE, but simply provides a reliable means to promptly, but also fallibly, distinguish between health and pathological conditions, it is still compatible with the possibility that the concept of DISEASE would be definable in terms of necessary (and sufficient) conditions.

In this sense, we shall maintain that the prototypical theory of disease, as developed, for instance, by Lilienfield and Marino (1995, 1999) or Sadeh-Zadeh (2000, 2008, 2011), can eventually be considered compatible with the "classical" conception of disease, which rather aims to find necessary and sufficient conditions to define the concept of DISEASE. Here the proposals are quite different to each other. It has been variously argued that all and only diseases are: biological dysfunctions, that is, conditions that are either impairments of normal functional ability (i.e., reductions of functional abilities below typical efficiency) or limitations on functional ability, caused by environmental agents (Boorse 1977, 1997, 2014); harmful dysfunctions, that is, conditions that result from the inability of some internal mechanism to perform that particular effect which is part of the evolutionary explanation of its existence and structure, and also cause some harm to the organism, as judged by cultural or social standards (Wakefield 1992); conditions that undermine one's own ability to fulfil all one's vital goals, that is those goals which are necessary and jointly sufficient for minimal happiness (Nordenfelt 1987); abnormal conditions that harm standard members of a species in standard circumstances, and for which medical intervention is both necessary and appropriate (Reznek 1987); conditions that are bad things to have, make the afflicted person unlucky, and are potentially medically treatable (Cooper 2002); incapacitating conditions that reduce one's own flourishing (Megone 2000, 2007).

To be clear, in the present paper we do not mean to take a side either in favour or against any of the above definitions of disease, nor to maintain that a classical and "conceptually-clean" definition of disease can eventually be found; less ambitiously, we would like to show a viable way for making the above classical characterisations compatible with the idea that the concept of DISEASE also has a prototypical structure. To put it another way, we would like to say that, from the point of view of the cognitive sciences, what is at stake are two different ways to represent the concept of DISEASE, produced by different cognitive systems, and these two kinds of representations are both important in their own specific domain and for their own particular scope. This means that, even if it would be possible to define the concept of DISEASE in classical terms through necessary and sufficient conditions, typical knowledge linked to the prototypical concept would still play a central role in many cognitive tasks, more precisely, in practical and clinical decision making. As we have shown in the previous section, we can account for typicality effects for classical concepts by adopting a distinction similar to the core/identification procedures distinction proposed by Miller and Johnson-Laird (1976): even if we have a classical definition of the concept of DISEASE (its core), in many practical cases it cannot be applied and, thus, we must trust fallible criteria coming from the prototypical concept of DISEASE.

At this stage, it is important to focus on the prototypical view about the concept of DISEASE in more detail. As we have seen in the previous sections, prototypical concepts are characterised by a set of features; none of these features is individually necessary and no group of them is jointly sufficient to determine category membership. In this sense, the prototypical disease is a theoretical entity that does not need to correspond to any real disease. Consequently, different conditions may still belong to the disease category even if they share just a few features or, at least in principle, no feature at all (providing that those features do constitute the disease category). That being said, the prototypical concept of DISEASE would obviously consist of both central and marginal examples. On the one hand, conditions such as tuberculosis, pneumonia or breast cancer are relatively prototypical examples for most individuals, and, thus, there is widespread agreement in considering them diseases. On the other hand, micromastia**,** cellulite or alcoholism may easily count as dubious conditions for many individuals, and, thus, there may be intractable disagreements regarding which of them truly belong to the category of diseases.

However, what might be the features constituting the category of disease that help us to promptly distinguish between health and pathological conditions? According to Merriam Webster, a disease is "a condition that prevents the body or mind from working normally; a problem that a person, group, organisation, or society has and cannot stop". Cambridge Dictionary similarly states that a disease is "something that is considered very bad in people or society". Another common view associates the concept of DISEASE with the lack of health, where health is defined by the World Health Organization as "a state of complete physical, mental and social well-being" (WHO, 1948). The above characterisations can be reasonably helpful to extrapolate some features commonly associated to the concept of disease; it is:

- a deviation from normality,
- not eliminable with a mere act of will,
- very bad,
- something that causes harm,
- something that affects one's own general well-being,
- a decrease of one's own physical functioning,
- a decrease of one's own psychological functioning,
- something that compromises one's own social role.

Looking at the philosophical literature where the prototypical conception of disease is proposed (see e.g. Lilienfield and Marino 1995, 1997), we may want to add other features such as:

- an evolutionary dysfunction;
- maladaptive;
- something with an onset, a course, and an outcome;
- medically treatable;
- something that obstructs the fulfilment of one's own goals.

Let's take stock here and assume that the prototypical concept of DISEASE would be constituted by all these features and, obviously, others. If someone does not associate many of the above features to the concept of DISEASE, we should probably conclude that they do not grasp the concept of DISEASE at all. This rough characterisation of the prototypical concept of DISEASE would assist us in making quick and reliable categorisations about diseases, which, in most cases, would be helpful to reliably and promptly distinguish between healthy and pathological conditions. That being said, the prototypical characterisation of disease may also lead to incorrect categorisations from the point of view of theoretical medicine, as all the above features are neither necessary nor sufficient conditions for being a disease.

For reasons of simplicity, let us focus on the prototypical structure of some syndrome-based descriptions of individual diseases. For example, influenza may be characterised by the following signs and symptoms (of course, among various others), which often guide everyday judgements and prompt clinical diagnoses:

- high fever
- coldness
- cough
- nasal congestion
- runny nose
- sneezing
- body aches

- sore throat
- fatigue
- headache

All the above features are signs and symptoms that characterise influenza in prototypical cases; however, none of them is either a necessary or a sufficient condition for correctly diagnosing influenza. On the one hand, the overall syndrome characterising the prototypical concept of INFLUENZA is clearly common to many other diseases, such as bacterial pneumonia (which, as the name suggests, obviously has a very different aetiology), and may thus lead to incorrect categorisations. Influenza is, in fact, an infectious disease caused by one of the three influenza viruses (*Influenza A virus, Influenza B virus,* or *Influenza C virus*), which thus count as necessary conditions for having influenza.

That being said, as we have argued in the previous section, associating prototypical information to concepts turns out to be useful in order to make defeasible inferences starting from incomplete knowledge, as in the case of medical diagnoses. In this respect, syndrome-based diagnoses are pragmatically useful and can be interpreted as specific-domain examples of the strong tendency of human subjects to resort to prototypical information in categorisation, even when it is not fully appropriate (at least from the point of view of theoretical medicine). To put it differently, even if we are able to define the concept of INFLUENZA in classical terms, at least through its necessary conditions (the presence of one of the three influenza viruses), typical knowledge based on the prototypical concept of INFLUENZA still plays a central role in many cognitive tasks, and, in particular, in doing quick and reliable diagnoses. For instance, as influenza reaches its peak prevalence in cold seasons and it is more prevalent than bacterial pneumonia, it would be reasonable to prescribe medications such as acetaminophen, instead of antibiotics, to an individual showing the above symptoms in winter. The use of typical knowledge in diagnostic tasks has to do with the constraints that concern every finite agent that has a limited access to the knowledge relevant for correctly categorising an individual disease.

4. Tentative conclusions

To sum up, disease concepts (both the concepts of individual diseases and the general concept of DISEASE) exhibit typicality effects that can be successfully explained in terms of prototypes. The prototypical characterisation of disease may lead to wrong characterisations from the point of view of theoretical medicine, however, it is still very useful to sort out pragmatic and clinical tasks in a prompt and reliable way. This means that the prototypical and the classical characterisations of the general concept of disease (and of the concepts of individual diseases as well) might still be compatible, as they respond to different cognitive tasks and are performed by different cognitive systems.

References

Armstrong, S.L., Gleitman, L.R. & Gleitman, H. (1983). What some concepts might not be, *Cognition*, 13: 293-308.

Blouw, P., Solodkin, E., Thagard, P., & Eliasmith, C. (2016). Concepts as semantic pointers: a framework and computational model. Cognitive science. no. 40:1128–1162

Borghini, "*De Principiis Generum Naturalium*", talk presented at the II FINO/SIFA Graduate Conference in Mind, Language, and Science "Between natural, social, and artificial kinds", University of Genova, June 30 and July 1, 2016.

Boorse, C. 1977. “Heath as a theoretical concept.” *Philosophy of Science* no. 44:542-573.

Boorse, C. 1997. “A rebuttal on health.” In *What is disease?*, edited by J.M. Humber and R.F. Almeder, 1-134. Totowa: Humana Press.

Boorse, C. 2014. “A Second Rebuttal On Health.” *Journal of Medicine and Philosophy* no. 39:683-724.

Cooper, R. 2002. “Disease.” *Studies in History and Philosophy of Biomedical Sciences* no. 33:263-282.

Frixione, M. (2013). Concepts and Fat Plants: Non-Classical Categories, Typicality Effects, Ecological Constraints, *Concepts – Contemporary and Historical Perspectives*, *ProtoSociology*, vol. 30, 152-166.

Gärdenfors, P., & Williams, M. A. (2001, August). Reasoning about categories in conceptual spaces. In IJCAI (pp. 385-392).

Giordano, L., Gliozzi, V., Olivetti, N., & Pozzato, G. L. (2013). A non-monotonic description logic for reasoning about typicality. *Artificial Intelligence*, 195, 165-202.

Harris, M. (1985). *Good to Eat. Riddles of Food and Culture*, Simon and Schuster, New York

Horty, J. F., Thomason, R. H., & Touretzky, D. S. (1990). A skeptical theory of inheritance in nonmonotonic semantic networks. *Artificial intelligence*, 42(2-3), 311-348.

Koons, R.C. (2013). Defeasible reasoning (revised version), Stanford Encyclopaedia of Philosophy (http://plato.stanford.edu/).

Lieto, A., Minieri, A., Piana, A., & Radicioni, D. P. (2015). A knowledge-based system for prototypical reasoning. *Connection Science*, 27(2):137-152. http://dx.doi.org/10.1080/09540091.2014.956292

Lilienfeld, S.O., and L. Marino. 1995. “Mental disorder as a Roschian concept: A critique of Wakefield’s “harmful dysfunction” analysis.” *Journal of Abnormal Psychology* no. 104 (3): 411-420.

Lilienfeld, S.O., and L. Marino. 1999. “Essentialism revisited: Evolutionary theory and the concept of mental disorder.” *Journal of Abnormal Psychology* no. 108 (3):400-411.

Medin, D.L., Wattenmaker, W.D. & Hampson, S.E. (1987). Family resemblance, conceptual cohesiveness, and category construction, *Cognitive Psychology*, 19, 242-279.

Megone, C. 2000. “Mental illness, human function, and values.” *Philosophy, Psychiatry, and Psychology* no. 7:45-65.

Megone, C. 2007. “Mental Illness, Metaphysics, Facts and Values.” *Philosophical Papers* no. 36:399-426.

Miller, G.A. & Johnson-Laird, Ph.N. (1976). *Language and Perception*, Harvard University Press, Cambridge, MA.

Minsky M., 1975. “A framework for representing knowledge”, in Patrick Winston (ed.), *The Psychology of Computer Vision,* McGraw-Hill, New York.

Nordenfelt, L. 1987. *On the nature of health: an action-theoretic approach*. 2nd edition 1995. Dordrecht: Reidel.

Putnam, H. (1975). *Mind, Language and Reality*, Cambridge University Press, Cambridge.

Rey, G. (1983). Concepts and stereotypes, *Cognition*, 15: 237-262.

Reznek, L. 1987. *The nature of disease*. London: Routledge & Kegan Paul.

Rosch, Eleonor, 1975. "Cognitive representation of semantic categories", Journal of Experimental Psychology, 104: 573–605.

Rosch, Eleonor, C. B. Mervis, 1975. "Family resemblances: Studies in the internal structure of categories", Cognitive psychology, 7 (4): 573-605.

Sadegh-Zadeh, K. 2000. "Fuzzy Health, Illness, and Disease." *Journal of Medicine and Philosophy* no. 25 (5):605-638.

Sadegh-Zadeh, K. 2008. "The Prototype Resemblance Theory of Disease." *Journal of Medicine and Philosophy* no. 33 (2):106-139.

Sadegh-Zadeh, K. 2011. *Handbook of analytic philosophy of medicine*. New York: Springer.

Socher, R., Chen, D., Manning, C. D., & Ng, A. (2013). Reasoning with neural tensor networks for knowledge base completion. In *Advances in Neural Information Processing Systems* (pp. 926-934).

Tversky, A. & Kahneman, D. (1983). Extension versus intuitive reasoning: The conjunction fallacy in probability judgment". *Psychological Review*. 90 (4): 293–315.

Wakefield, J.C. 1992. "The concept of mental disorder. On the boundary between biological facts and social values." *American Psychologist* no. 47:373-388.

World Health Organization. "Constitution." By Arthur L. Caplan, H. Tristram Engelhardt and James J. McCartney. *Official Records of the World Health Organization* 2 (1948): 100.