One Solved and One Unsolved Problem for Conceptual Atomism

Manuel Bremer
Universität Düsseldorf, Germany
www.mbph.de

In this talk I consider two problems for conceptual atomism. Conceptual atomism can be defended against the criticism that it seems to contend that all concepts are simply innate (even technical concepts to pre-technological humanoids) by specifying the innateness thesis as one of mechanisms of hooking up mental representations (concepts as *language of thought* types) to properties in the world (§1). This theory faces a problem with non-referring expressions/concepts, it seems. Conceptual atomism can, however, deal with non-referring expressions/concepts (§2). Hooking up concepts with properties raises, further on, broader metaphysical problems of making concepts correspond to (natural) properties. These questions are much harder to answer (§3).

§1.

Conceptual atomism (CA) is the claim that many if not most concepts cannot be decomposed into a set of conceptual parts or features thus that this set of features is not just necessary but also provides a sufficient analysis of the concept, thus that the conjunction of the features is equivalent to the concept in question. Of course there are lots of concepts that are derived compositionally from these atomic concepts, and these derived concepts can, obviously, be decomposed again. Still the majority of concepts – and especially those concepts that are in other semantic theories considered to be decomposable – are not decomposable to the conceptual atomist. I will not rehearse the arguments for conceptual atomism here (cf. Fodor 1998, 1998a), Fodor's main arguments for CA are mostly negative anyway: other theories do not deliver what they promise: there are neither enough good definitions around for theories of lexical decomposition nor can theories of prototypicality provide compositionality. CA by definition does not bother about the absence of definitions, and it delivers compositionality for complex concepts, the existence of which is, of course, not denied, and for sentences as well.

The main criticism of CA mobilizes some version of the 'incredulous stare' that most concepts (even REFRIGERATOR) should be atomic and innate. This, however, mellows down to an over-interpretation and some simplistic ways of expressing CA (for the purpose of getting attention …). The point, of course, cannot be that REFRIGERATOR itself is innate (as some *language of thought*-type [cf. Fodor 1975, 1987]) but the means of atomistically tying up to refrigerators. In fact artefacts might be a bad example, since entities or corresponding theoretical concepts that are explicitly defined are not excluded by CA. CA just claims that

more concepts are atomic than one thought of before. DOG might be a better case in point. Discussing the concept of DOORKNOB Fodor (1998: 136-37) himself stresses that having some concept means *resonating to some property* in the appropriate fashion. Even if the physical structure DOORKNOB is composite the concept *might* be unanalysable into jointly sufficient semantic features. Therefore DOORKNOB might be primitive to us, but this does not mean that the LoT-symbol for DOORKNOB itself is innate. What are innate are the cognitive path ways and mechanisms that introduce us to the concept DOORKNOB in the presence of DOORKNOB:

> all that needs to be innate is the sensorium. … [T]he kind of nativism about DOORKNOB that an informational atomist has to put up with is perhaps not one of *concepts* but of *mechanisms*. (ibid p.142).

That sounds far less eccentric than Assyrians running around with ACCELARATOR.

That a concept cannot be defined and thus has to be taken as atomic and thus – at least in principle – as isolated from supposedly closely related concepts – related in which way, anyhow? – does not mean that in the normal process of acquisition a concept comes alone. Given objective relations (i.e. relations pertaining in reality) between the properties referred to, and given the probable sensitivities of the cognitive mechanisms which link a concept to its referent, we may suppose that having these mechanisms in place and interacting with an environment which comes with a structure results in acquiring concepts more or the less at the same time, or, at least, acquiring a concept and a couple of related concepts referring to (metaphysically or nomologically) related properties. Thus DOOR may be not a 'feature' present because of a complete decomposition of DOORKNOB, but DOOR and DOOR may be – under normal circumstances – be co-present with DOORKNOB and DOORKNOB, because doorknobs are – under normal circumstances – on doors. A mind/brain able to link to doorknobs is likely to be linked to doors as well. And so for many other concepts.

The co-presence of DOOR and DOORKNOB can be explained this way. It does not cut in favour of some semantic theory. A question arising with this, however, is what else one may expect to be present once both concepts are present in a conceptual system. There should be some analytical links between them.

In contrast to DOOR and DOORKNOB it is not beyond imagination that one can hook up – for some time at least – to SCAR without having WOUND, which is unfavourable to a semantic analysis involving historic knowledge for some concepts.

§2
Non-referring terms pose a *prima facie* problem for any informational or externalist account: If there is nothing, there is no referent, there is no content. So there is no concept. But we have a lot of non-referring terms. Thus the problem for CA. There are several kinds of non-referring terms to distinguish:

(2.1)   Descriptive terms which fail to refer

(2.2)   Entities presented as fictional (fiction presented as fiction)

(2.3)   Entities presented as real which do not exist (fiction presented as real)

(ad 2.1)

Failing descriptive terms pose no problem. Descriptions are compositional. Speaking of "the third student that handed in a paper on February 2nd 2008" and then checking the behaviour of students at the contextually presupposed bureau one finds out that less than three students handed in a paper on that day, so that there was no third student handing in a paper, so that the description does not pick out an entity. Since all the words making up the description have a meaning (the concepts they express have content) there is no problem for CA here.

(ad 2.2)

A theory of fiction raises a lot of difficult questions about modality, withholding epistemic commitment, and contextual understanding I cannot go into here. In as much as referring is an intention a speaker associates with some of the expressions she uses, presenting fiction as fiction will not involve this intention. One may ask, still, about the content of the expression for fictional entities. It seems that bracketing our commitment to the reality of the described or depicted events sets into motion another mechanism for acquiring labels and concepts. The sensory capacities of our mind/brain work to hook us up to real properties. The imaginary capacities of our mind/brain may allow introducing concepts (like the kind VAMPIRE) and hooking them up not by a reliable connection to some part of reality but by a loose connection to a file of information provided by the fictional context. VAMPIRE need not come to be defined by this – if fictional entities are defined they are just another instance of (2.1) – VAMPIRE refers to the vampires as being presented as real in the file of fiction. Many fictional entities are explicitly defined – often similar to theoretical terms if you think of science fiction – and thus reduce to problem (2.1). Some are developed as the story or movie enfolds.

As we as readers also have access to our beliefs and some knowledge of our locking in to the world we can stabilize such a concept by claiming some counterfactuals like "If I was in that world and in *that* situation I would see a vampire". VAMPIRE then is (roughly) synonymous to PROPERTY THAT A HUMAN RECOGNIZED IN STOKER'S WORLD IN THE PRESENCE OF VAMPIRES, where "VAMPIRE" is – main option – understood as being indexed to the world as presented by Stoker. Second option: If there metaphysically is – due to the existence of other possible worlds – the property VAMPIRE, there is no problem about the expression "vampire" and VAMPIRE in the first place, since they just refer to it, even if reference now is – strangely enough – mediated by fiction as a channel.

Things get more involved if the fiction also pertains to the cognitive systems and their capacities. But given a good enough theory of fiction – I do not have – cases (2.2) pose no problem.

(ad 2.3)

The hard case of non-reference seems to be that some expression or concept is used *prima facie* and intentionally as referring, but never had any reference. If it never had had any reference, how could it have been used or understood in terms of CA?

A first simple solution appeals to making mistakes. Our mechanisms of acquiring concepts malfunctioned in this case. Some supposed clues to a supposed referent were introduced, but corresponded to nothing (i.e. not to a unique property). Beliefs and statements involving that concept or term were false (or meaningless). End of story, even reliable mechanism allow for the occasional failure.

One may reply that there are quite a lot of non-referring expressions. So we do not have just an occasional failure. On the other hand this reply had to show that enough of the non-referring cases belong outside of categories (2.1) and (2.2).

A second attempt at a solution may try to assimilate the cases of (2.3) to the cases (2.1) and (2.2). For instance, the concept ETHER of 19th century physics either was introduced as a defined theoretical term (reduction to case 2.1) or one may have understood ETHER in terms of counterfactual encounters with ETHER (reduction to case 2.2).

A third attempt at a solution may try to use the limited truth of use theories of meaning. Referential content is tied to concepts. We often meet the problem of non-referring expressions with respect to words. Usually a word may be simply linked to a concept and thus we have the harder problem of non-referring concepts. On the other hand, word-concept marriages are forged and maintained by conventions of usage. A convention lives as a stable regularity in a community of language users. It does not depend – at least in the short run – on a word having objective content. At least for a while, which might last well dozens of years, a term may be in use on the erroneous presumption of a common understanding. This is an observation for a theory of conventions, not a problem for semantics.

A rejoinder to this third option is that apart from the problem of non-referring expressions there is the problem of non-referring *concepts* (LoT-tokens) since we can *think* about non-existents, and the medium of thought is the *language of thought* (i.e. not the public language). Once again, however, the third option has decreased the remaining critical cases: still critical are 'only' cases where neither descriptions are involved, nor fictional contexts introducing the concept, nor linguistic items standing in for supposedly non-referring concepts. One may also add to the third option the idea that we have to allow for concepts to refer to linguistic items, since after all we also think *about* language. Then we may allow for the case that a concept (thus a vehicle of thought) is present, but the concept is a concept like WHATEVER IS REFERRED

TO BY "ether", so that the concept is partially *meta-linguistic*. This is a non-referring concept, but none that poses a problem for CA.

The ultimate attempt at a solution is the introduction of *possibilia*. *Possibilia* which do not exist can stand in as referents for any problematic concept. For informational semantics coupled with this dose of modal realism one still has well-defined infons containing non-existing *possibilia* along with existents (i.e. existing *possibilia*). This, of course, is a heavy ontological investment, but the problem of *possibilia* may turn up anyway in the broader metaphysical picture of things.

§3

Even if concepts are either clear as they are atomic or as they are defined concepts are not lexical items. Already the naming conventions for concepts suggest the idea of one property: one concept. In proportion to the limited sensibilities of the mechanisms that link concepts to their referents there is *some* indeterminacy of reference (say if the mechanism cannot distinguish between some nomologically co-instantiated properties, and one's property theory counts them still as two properties). The larger problem might be shifted to lexical items.

A lexical item – in the ideal case – has one concept as its meaning.

Nothing excludes, however, that one lexical item has more than one concept as its meaning. In cases of clear ambiguity the lexical item can be treated as ambiguous, with the two readings known and separable; there could as well be two lexical entries in this case.

In other cases there may be distinct but in their content (i.e. the properties referred to) overlapping concepts, where we do not recognize that we are dealing with two concepts. In many applications of the lexical item the differences may not count or surface, but in others they may well do so.

A – so to say – well-behaved realm of concepts does not guarantee a well-behaved lexicon.

How could the cognitive mechanism that links concepts to linguistic expressions work? The most simple idea might look like a stamp engine:
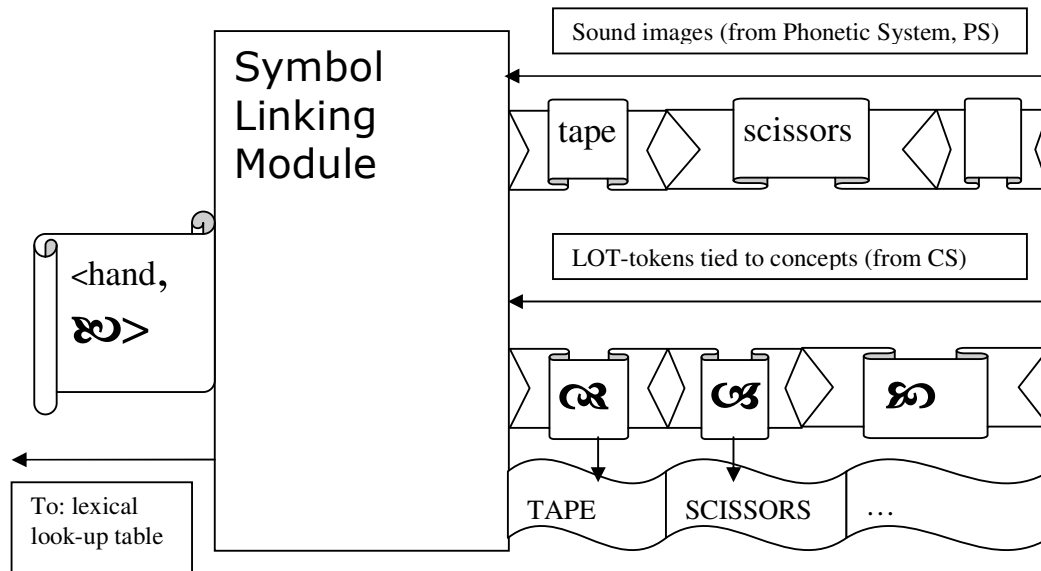
NAÏVE PICTURE OF SYMBOL LINKING

There is an input belt to the labelling box; in each cell of the input bell sits a token of a LoT symbol type; this token is labelled with a token of a linguistic expression and the pair is stored into a look-up table. Now, if a token of the LoT type is to be expressed or an expression to be decoded the look-up table is consulted. If that was the case the ideal of an injective mapping between LoT and language (and a bijection between coded LoT types and lexical items) would hold. Ambiguity would only occur because of malfunctioning: either stamping two LoT tokens at the same time or multiply assigning a lexical item to several LoT tokens. How the labelling mechanism works may be considered as *the soft symbol link problem*. Dealing with it one has to provide a theory that combines an account of the mechanism as just outlined from the perspective of the *individual* working cognitive system with an account of conventions and/or coordinative functions, which play a major role in selecting some word (type) as expressing a concept within a language *community*. This problem includes some tricky questions of social coordination, it seems, but may well be put from the perspective of the intersecting idiolects of individual speakers (i.e. be put from the individual's perspective). In any case the *lexical* assignment problem might be considered to be not that crucial for a theory *of concepts*.

Things – in the most literal sense – may not be that simple. There are properties out there. These properties stand in relations, some of which are mereological. Thus properties can be combined, whether this combination is natural in the sense of spatial or causal contiguity or connection or not. The thus built complex is metaphysically real, simply said: a property itself. Thus in the absence of any guarantee that our cognition-world alignment mechanism

6

are tuned in to natural properties (in a sense of "natural" to be specified in a substantial metaphysical theory) we may have tuned in to such a property and so have developed a concept that as every concept traces some property out there, but trace a property that as well could be spilt up in two more natural properties (again in a substantial sense to be specified). In this case the conundrums presented by some lexical items could go back to such gerrymandered concepts. Can we ever distinguish a case where two clear concepts are tied to one lexical item from a case where a gerrymandered concept got lexicalized? And is that an important distinction in the first place? If it is not an important distinction in the first place, what does this tell us about the benefits of a well-behaved conceptual system? The plethora of these questions may be considered *the hard symbol link problem.*

One may think that concepts are individuated by the properties they are hooked up to. But here lays some metaphysical problem: The hooking-up uses *nomological* mechanisms. It is tied up to the laws of nature as they are. Properties on the other hand – at least in many, typically categorical, non-dispositional theories of them – are intensional entities. Even if some other laws *governed* the behaviour of a property it still would *be* the same property. That is the mark of categorical property theories (like Armstrong 1978, 1997). The rationale behind that conception is that laws are higher order facts relating properties ('constraints' in situation semantics talk). These facts should be independent from the first order facts, which object exhibits which (basic) properties.

Now, what counterfactuals does a LoT-symbol α have to support to count as a representation of the property *F*? Nomological counterfactuals may not be enough if properties are intensional. Modal counterfactuals α cannot support by being part of a cognitive system built according to the laws of nature as they are in our world. Should we commit us with CA to dispositionalist property theories, which identify a property with its causal powers (cf. Shoemaker 1984), then?

One need not do that if one links CA to some version of *two-dimensional semantics* (cf. Chalmers 1996: 56-71). According to two-dimensionalism one might say: The concept α is the concept it is, because it reliably (in terms of the existing laws of nature) refers to some property, which can even by presented *by that very concept* α as being present in some other possible world with *different* laws of nature. Concepts thus defined underwrite modal talk. WATER is hooked up by some identificatory mechanisms with instances of WATER. These known or unknown mechanisms constitute the A-intension of "water" (the term which expresses WATER) – or they constitute at least a part of it as we have no full definitions in CA. And "water" and WATER refer in all possible worlds by their B-intension to the same stuff, namely the stuff they pick out in our world: WATER.

One problem remains: If properties can be separated from their nomological behaviour then the fact that two properties are nomologically co-instantiated in our world does not make them one property. In other possible worlds they are governed by different natural laws, and thus are not co-instantiated. The resulting problem for the option outlined in the paragraph before then is: Which of the two is the referent of the concept? The A-intension of the concept

picks out both of them, but they are not the same stuff. One alternative is to allow for disjunctive properties, thus concepts still being in 1:1-correspondence with properties. Another alternative is to admit some limits of our referential mechanisms, as they came about in our world and thus by nomological necessity are blind to the nomological identities of our world. So for some concepts we cannot individuate a unique referent.*

References

Armstrong, D.M. (1978). *Universals and Scientific Realism.* 2 Volumes, Cambridge.
      - (1997). *A World of States of Affairs.* Cambridge.
Chalmers, David (1996). *The Conscious Mind.* Oxford.
Cherniak, Christopher (1986). *Minimal Rationality.* Cambridge/Ma, 2nd Ed.
Fodor, Jerry (1975). *The Language of Thought.* New York.
      - (1987). *Psychosemantics.* Cambridge/MA.
      - (1998). *Concepts.* Where Cognitive Science Went Wrong. Oxford.
      - (1998a). *In Critical Condition.* Polemical Essays on Cognitive Science and the Philosophy of Mind. Cambridge/MA.
Shoemaker, Sydney (1984). *Identity, Cause and Mind.* Philosophical Essays. London.

---

*   The material presented is taken from my book *Conceptual Atomism and Justificationist Semantics.* Bern et al. (Peter Lang Publishers), 2008.