Knowledge Relativity

Oliver Hoffmann oliver@hoffmann.org

University of South Australia

Abstract. I introduce the notion of *knowledge relativity* as a proposed conceptual link between different scientific disciplines. Examples from Informatics and Philosophy, particularly Newell's *knowledge level* hypothesis and Popper's *world 3* of *knowledge*, are used to demonstrate the motivation for making this notion explicit.

1 Why Knowledge Relativity?

Not every assumption is *knowledge* and whether somebody's personal views on the world are classified as *knowledge* is determined by complicated interactions among human subjects and between human subjects and their environment.

The word *science* itself means *knowledge* and thus one would assume that the distinction between arbitrary unilateral assumptions on the one hand and shared mutually assured *knowledge* on the other hand is uncontested in contemporary science. But one scientific discipline has adopted a notion of *knowledge* that abandons the distinction between *knowledge* itself on the one hand and the mode of it's expression or temporary attempts to arrive at it on the other hand: In Informatics, *knowledge representation* is progressively equated with *knowledge* itself, and isolated pieces of *knowledge representation* stored in a single computer system with mutually confirmed *truth*. Philosophy would be ideally situated to contribute to a more mature notion of *knowledge* in Informatics, but the relativity of *knowledge* appears to be too commonplace in contemporary Philosophy to be explicitly stated.

In this article, I will therefore try to explicate *knowledge relativity* and the potential role of this notion both in Informatics and in Philosophy. The starting point for this task will be asking the missing "Who" question for some of the basic concepts in Informatics and the result will be a notion of *knowledge* that emphasizes *subjects* over *objects* and involves an understanding of *standardization* instead of the predominant notion of *truth*.

2 The Knowledge Level

Since it's early beginnings, Informatics as a scientific field has favored concepts and models that allow for the marginalization of subjectivity and relativistic views. One of the numerous examples for this tendency is Newell's hypothesis of a *knowledge level* as the sole and exhaustive location for *knowledge* in a system with artificial intelligence (Newell 1982). This level would have all the properties of a normal computer system level: among other properties, it can be implemented without references to internal details of other levels, and it can be reduced to the level below it (the *symbol level*) by defining it's medium (*knowledge*), components and laws via those (*symbols*) of the level below it. According to this hypothesis, *knowledge* could be seen as an abstract property of (computer system or human) *agents* implementable via various different forms of symbolic representations in the same way that symbolic representations are implementable via various different forms of electronic hardware. Such an understanding of *knowledge* would make it impossible to directly verify or falsify the presence of a specific element of *knowledge* in an artificial *agent*: the highest directly observable system level contains symbols that might or might not encode a specific *knowledge*, but the *knowledge* itself would reside one level above those representations and would be removed from direct observation. But Newell proposed a mechanism of indirect verification for the *knowledge level*: If some (human or artificial) *agent* A can detect the impact of some specific *knowledge* in the actions of another *agent* B, *agent* A can verify the presence of such *knowledge* in *agent* B (Fig. 1). Through this hypothesis, Newell

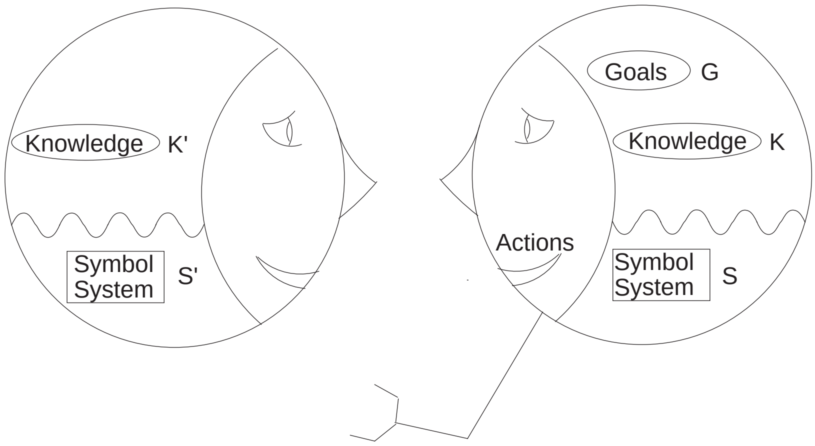

Fig. 1. verification of knowledge according to Newell

attempted to end an internal dispute among artificial intelligence researchers, a controversy centered around the question of the "best" form of *knowledge representation*. According to Newell himself (Newell 1993), this attempt was partially successful, and a significant part of artificial intelligence research has been (implicitly or explicitly) based on the *knowledge level hypothesis*. I will not discuss the correctness of Newell's hypothesis in this article, partially because it is too complex to determine how it could be verified or falsified at all, but also because the correctness or incorrectness of this hypothesis is not directly linked to the central argument in this article: the general tendency of Informatics to suppress notions of relativity and the need to introduce explicit terminology for expressing relativity. A starting point for detecting the implicit *objectivist* world view (Stamper 1993) underlying artificial intelligence concepts like the *knowledge level* is to focus on those aspects that authors like Newell do not discuss,

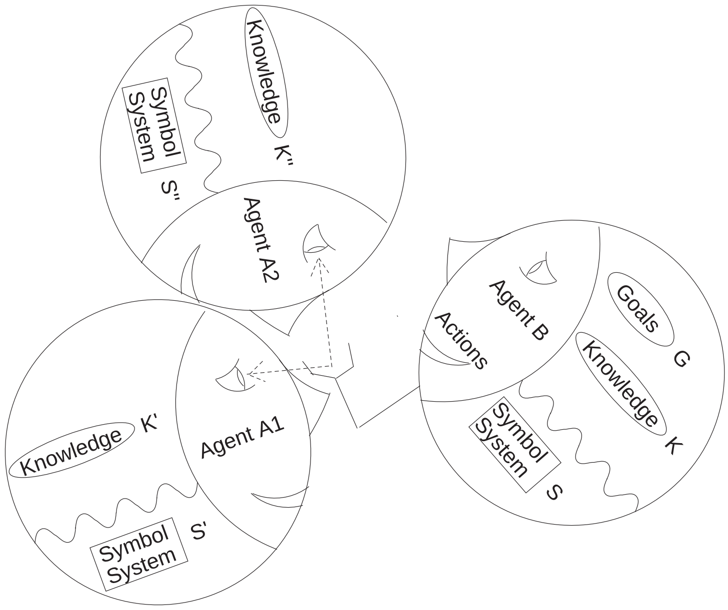

Fig. 2. how do different observing agents agree on the presence of knowledge?

and the questions that are not answered. A good example is: ”Who is the *agent* that will verify the presence of *knowledge* in another *agent*?” More than one *agent* could take the role of the observing agent A in Newell’s *knowledge* verification mechanism and the different As could come to different conclusions on whether *knowledge* is present in *agent* B (Fig. 2). Newell does not explain how the A *agents* would be able to come to any consistent conclusion on *agent* B’s knowledge, and since their symbol levels might be mutually incompatible, it is unclear how they could even communicate their different views on *agent* B. The only conceivable way Newell’s *knowledge* verification might yield consistent results is by relying on a standardized observer *agent* A that serves as the absolute reference on detecting *knowledge*. Newell implicitly assumes *objective knowledge* that would be the basis, rather than the result of attempts to create artificial intelligence. But if such a standard *agent* existed, Newell’s entire *knowledge* verification procedure might be obsolete before it is ever applied: this standard verification *agent* would already incorporate the perfect embodiment of intelligence and *knowledge*, and any attempts to build more intelligent *agents* would have to fail. The main effect of Newell’s *knowledge level hypothesis* is not what it accomplishes, but what it prevents: the authority to interpret symbolic *knowledge representations* is restricted to computer systems, since their (virtual) behavior is the only way for determining what *knowledge* is represented. Human subjective interpretation is excluded from the process of symbol interpretation, since Newell’s artificial *agents* never directly expose their symbol level to human *agents*.

3 Information Theory Roots

Newell's approach of replacing human *subjective* interpretation with the implicit assumption of an *objectively* knowledgable observer represents the general trend in Informatics, and both method and motivation of this approach can be traced back to the origins of the discipline: The original technical definition of *information* (Shannon 1948) is based on the ability to predict the next symbol in a stream of communication. Again, the open question is: "Who will predict the next symbol?" Different potential receivers of the same symbolic message will have different *knowledge* about the world in general, the language used for sending the message, and the sender. Thus different receivers will have different abilities to predict the next symbol, and the *information* content of the same message will potentially be different for each receiver (Fig. 3). Shannon arrives at an *objectivist* result by assuming the existence of a *subject*-independent dictionary containing absolute probabilities for a given language. The receiver of Shannon's messages turns out to be the same implicitly standardized perfectly knowledgable observer that Newell uses to verify the existence of *knowledge* in computer systems. Shannon, like Newell, was motivated by the goal of eliminating subjective interpretation from his model of symbolic message content, and his *information* notion is still in use today. Informatics has suppressed *knowledge relativity* since it's beginning, and current trends to equate *knowledge* with *knowledge representation* are the direct continuation of this tradition. The discipline initially had good reasons for choosing this approach: At the time of Shannon, technical reliability of communication was the primary goal, and side effects on the potential processing of stored *knowledge* inside computer systems were largely irrelevant to the engineers that founded Informatics. Wherever issues of *knowledge relativity* come in conflict with engineering properties like reliability, predictability, and consistency, Informatics has generally given preference to those views that favor engineering goals. Implicitly, *knowledge relativity* has always played a role in Informatics: as the notion that has to be eliminated.

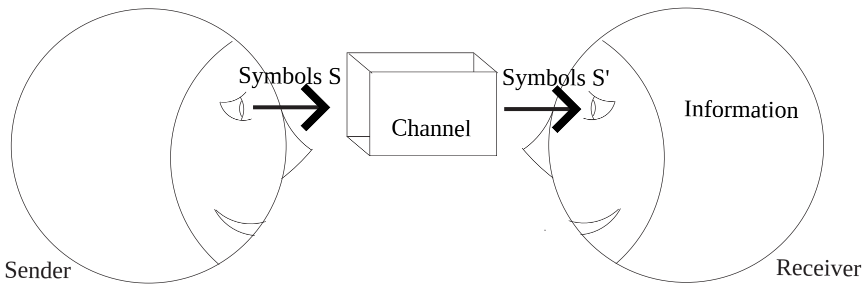

Fig. 3. different receiver *agents* would receive different amounts of Shannon *information*

4 Three Worlds

Due to it's longer history, Philosophy has witnessed more changes in the role of *knowledge relativity* than Informatics. During it's co-existence with Informatics in the last 60 years, however, *knowledge relativity* played a comparatively implicit role in Philosophy, much like in Informatics. But in contrast to Informatics, *knowledge relativity* was implicitly treated as a given basis of inquiry. To illustrate main differences in the implicit

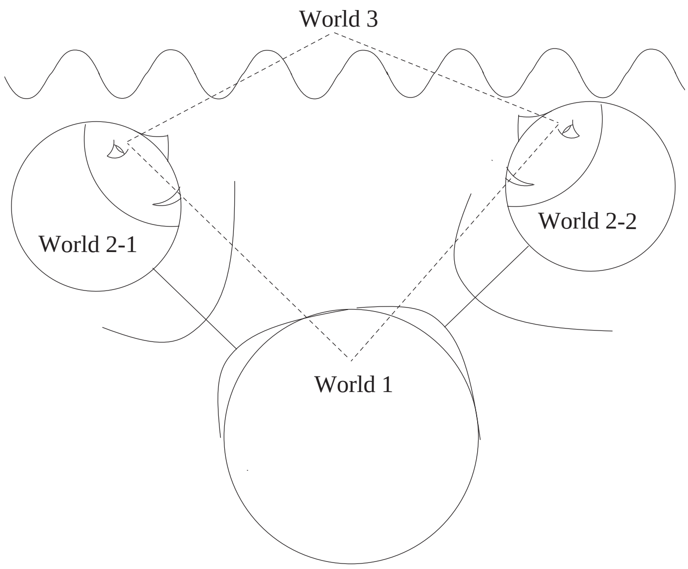

Fig. 4. Popper's *3 worlds* thesis

treatment of *knowledge relativity*, I will discuss a prominent example from the Philosophy of Science: Popper's *3 worlds* thesis (Popper 1972). Popper defines 3 different *worlds* that are interrelated but separate (Fig. 4): *World 1* is the (observer-independent) physical universe, *world 2* is an observer's image of *world 1*, and *world 3* is the collective symbolic *representation* of *world 2* shared among a multitude of observers. *World 3* is necessarily embedded in *world 1*, since observers are only able to communicate each other's symbolic *representations* if they are able to physically perceive them. Furthermore, any action performed by one of the observers can be interpreted as a contribution to *world 3*, and scientific experiments are the most prominent example of this feature. Popper's motivation for his *3 worlds* thesis was somewhat similar to Newell's motivation for his *knowledge level* hypothesis: Popper wanted to show a path towards *objective* observer-independent *knowledge*. But in contrast to predominant Informatics approaches, Popper assumed *subject*-relative *knowledge* and *subject*-relative *representation* as the starting point of this path. *Objective knowledge* is only achieved as the result of a long process of negotiation among *subjects*. This process is permanent, since *world 3* only approximates *world 1*, without ever reaching total consistency with it.

5 A Matter of Perspective

The reception of Philosophical concepts in Informatics is somewhat ambiguous: Informatics approaches frequently borrow their terminology - for instance *ontology* (Gruber 1991) - from philosophy, but later transform the meaning of these terms in very drastic ways, at times even reversing their original meaning. I argue that this phenomenon is

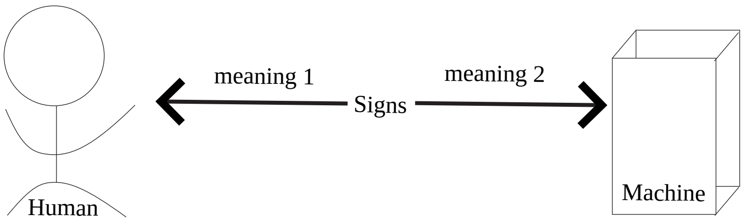

Fig. 5. dual role of the algorithmic sign

due to the dual role of symbolic *representations* in Informatics: During computer system programming, symbolic *representations* are created and interpreted by humans, but during computer system execution time, these *representations* are interpreted and modified by machines. The *algorithmic sign* (Nake and Grabowski 2001) (Fig. 5) incorporates a dual role towards the human observer on the one hand and the computer system on the other hand, at times as an external symbol used by human *agent*(s), and at times as an internal symbol used by machine(s). Machines are standardized products, and they are expected to work according to specifications. The same symbolic *representation* would therefore be expected to be open for re-interpretation when read and analyzed by different human *agents*, but would be expected to be stable and rigid in it's meaning and function when processed by different computer system *agents*. Since the final goal of Informatics is creating computer systems, the discipline has typically favored views that marginalize human *subject*-relative aspects. This approach has proven successful in areas that require predictability and consistency, but unsuccessful when the cost of implicit standardization was too high. One example for the latter are *knowledge* management tasks or the computer support of human creativity. For a more successful interaction with computer systems in these areas, *knowledge relativity* has to be made explicit.

6 Dimensions of Knowledge Relativity

In order to explicate the relativity of *knowledge* I will try to list the different dimensions of *knowledge relativity*:

1. *subject* who is holding the knowledge: who knows?
2. *representation* what kind of symbolic expression is used to communicate the knowledge: what was expressed?
3. *evaluator* who is judging the presence of knowledge: who thinks someone knows?
4. *communicative intent* was the *knowledge* expressed in order to inform or in order to reach agreement on known issues: do we need to discuss this?
5. *functional intent* what use was intended for the representation used: what will it change?
6. *receiver* who was the intended receiver for the *representation* used: who should know?

Not all of these dimensions need to be fully developed. A human agent might have *knowledge* without communicating it in any form, for instance, so the *representation* dimension would not be developed. Some dimensions might also have identical values, the *subject* holding the *knowledge* and the *evaluator* judging it's presence might for instance be identical ("I know"). The primary line of distinction in the treatment of *knowledge* between human communication and machine intelligence can be found in the *functional intent* (will the *representation* be interpreted as an external signal or directly processed inside the system?), and the primary line of distinction between different disciplines like Philosophy and Informatics can be found in the *communicative intent* (do we want to declare a standard or start a discussion?). Such a definition of

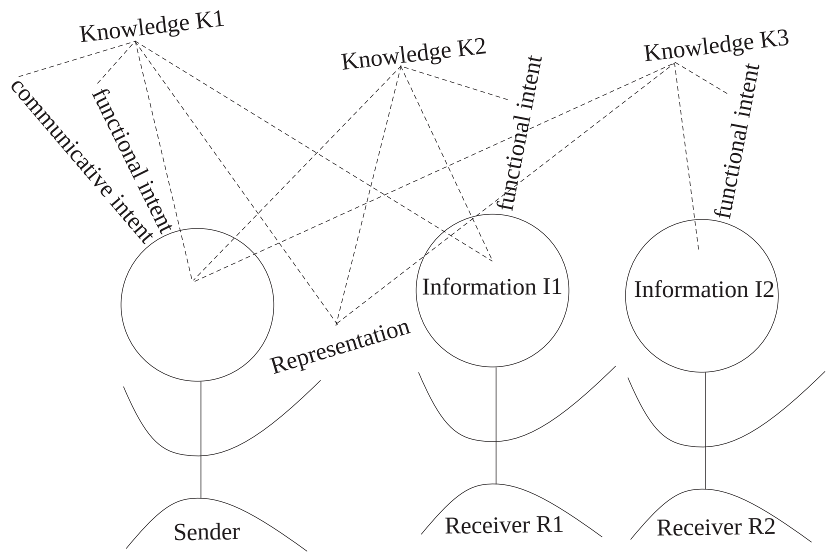

Fig. 6. *knowledge relativity* aware version of Shannon's *information*

knowledge relativity does not require reference to any *object*, but it requires reference to *subjects*, as can be demonstrated on the example of a *knowledge relativity* aware version of Shannon's *information* (Fig. 6): A signal (the *representation* is assumed to already contain any channel distortion, S' in Fig. 3) sent by one *subject* could be *received* by a multitude of *subjects*. For each *receiver*, this signal could *represent* different *knowledge*, always in relation to each *receiver's knowledge* about the *sender* and each *receiver's intentions* for this signal. Thus each *receiver* could detect a different amount of *information* for the same signal, and the same signal would contribute to the *knowledge* of different *receivers* in different ways. For an individual *receiver*, Shannon's *information* definition as the degree of un-ability for predicting the next symbol (in relation to this *receiver's knowledge*) still holds. From a *knowledge relativity* aware perspective, *knowledge* is not contained in any of the involved *subjects*, nor can it be assembled from or reduced to *subject*-external components like those constituting symbolic *representations*: Since any understanding of the link between some specific *subject*, some specific *representation* and some specific *knowledge* always requires some other *subject* in the role of *evaluator*, the link between *representation* and *knowledge* will always be dynamic.

7 Conclusion

The main motivation for the introduction of *knowledge relativity* in this article was to facilitate the dialogue across disciplines, particularly between Informatics and Philosophy. But assuming a fertile dialogue and assuming this new notion proves useful, I would expect a direct effect both in Informatics and in Philosophy, and potentially in more disciplines. A fair amount of literature outside Informatics and Philosophy deals with issues similar to the ones exemplified above, and if it is true that we already live in a "knowledge society", some additional clarity on the nature of knowledge or at least the nature of the term *knowledge* should prove useful.

References

Gruber, T. R.: 1991, The Role of Common Ontology in Achieving Sharable, Reusable Knowledge Bases, *Principles of Knowledge Representation and Reasoning*, Morgan Kaufmann, Cambridge, USA.

Nake, F. and Grabowski, S.: 2001, Human-computer interaction viewed as pseudo-communication, *Knowledge Based Systems* **14**.

Newell, A.: 1982, The Knowledge Level, *Artificial Intelligence* **18**, 87–127.

Newell, A.: 1993, Reflections on the knowledge level, *Artificial Intelligence* **59**, 31–38.

Popper, K.: 1972, *Objective Knowledge: An Evolutionary Approach*, Oxford University Press, UK.

Shannon, C. E.: 1948, A mathematical theory of communication, *Bell System Technical Journal* **27**.

Stamper, R. K.: 1993, Social norms in requirements analysis, *in* M. Jirotka, J. Goguen and M. Bickerton (eds), *Requirements Engineering, Technical and Social Aspects*, Academic Press, New York, USA.