Human Action in the Healthcare Domain: A Critical Analysis of HL7's Reference Information Model[*]

Barry Smith
Lowell Vizenor
Werner Ceusters

Abstract

If we are to develop efficient, reliable and secure means for sharing information across healthcare systems and organizations, then a careful analysis of human actions will be needed. To address this need, the HL7 organization has proposed its Reference Information Model (RIM), which is designed to provide a comprehensive representation of the entire domain of healthcare centered around the phenomenon of human action. Taking the Basic Formal Ontology as our starting point, we examine the RIM from an ontological point of view, describing how it fails to provide a representation of the healthcare domain which would enjoy the sort of clarity, coherence, rigor and completeness that is claimed on its behalf.

1. Introduction

Information and communication technology has not only altered the way that medical information is generated, stored, analyzed, and shared across and within healthcare organizations, it has also come to be associated with the promise that it will increase the safety, efficiency and cost-effectiveness of healthcare. The hope is that the electronic health record (EHR) and associated reporting, analysis and decision support technologies will facilitate the diffusion and dissemination of healthcare information, thereby allowing the systematic use of clinical guidelines

---

[*] We dedicate this paper to Ingvar Johansson our friend and colleague at the Institute for Formal Ontology and Medical Information Science (IFOMIS) from 2002 to 2008. The interdisciplinary research group at IFOMIS included representatives from ontology, biomedical informatics, and linguistics, and Ingvar himself contributed in important ways in all of these areas — including the peculiar interaction between ontology, medicine and speech act theory which forms our topic in what follows.

and outcomes measures in ways that will bring benefits to human health in the form of increased safety, effectiveness and economy.

It has for long been clear, however, that many difficulties must be overcome before the promise of health informatics can be fully harvested. One such difficulty arises from the fact that the single doctor — single patient nexus has been largely superseded by a regime in which the typical patient is managed by a team of health care professionals, each specializing in one aspect of care. This is significant because different departments within a healthcare organization have different disciplinary cultures and employ different terminologies and data formats to talk about what are putatively the same phenomena. Current efforts to develop efficient means for sharing information across healthcare systems and organizations that have some prospect of overcoming this and a range of similar problems must find effective ways to share information in an intuitive and stable way that ensures that meaning is preserved.

2. The Birth of the HL7 RIM

To see how difficult a task this is, we describe one ambitious and highly influential effort to standardize healthcare information across the entire domain of healthcare that has been advanced by the Health Level 7 (HL7) organization, one of several ANSI-accredited Standards Developing Organizations operating in the healthcare arena.[1]

Already in the 1990s HL7 enjoyed considerable success through its creation of a widely used standard for healthcare messaging which was established as mandatory for communication between US Federal Government-funded healthcare organizations. This standard is now commonly referred to as HL7 version 2 — or v2, for short.[2] v2 enabled healthcare applications to exchange clinical, demographic and administrative data in digital form on the basis of what we can think of as a walkie-talkie paradigm. Significantly, the v2 standard was designed to meet the interface requirements of the healthcare system in its entirety rather than focusing on the requirements of just one area of healthcare such as pharmacy, imaging services or insurance claims management.

---

[1] http://hl7.org, last accessed November 12, 2012.
[2] We here ignore the differences between successive sub-versions of the v2 standard.

Unfortunately, as v2 proved ever more popular, it also led to the creation of manifold v2 dialects, which over time brought about a situation in which messaging interoperability was maintained in many cases only within, and not between, healthcare organizations. Fatefully, the response of the HL7 organization to this problem was to develop, starting in the 1990s, an abstract model of the entire domain of patient care called the Reference Information Model (RIM) that was intended to serve as a unified framework for the sharing of information and the usage of data across the entire domain of healthcare and to serve as a constraint on all subsequent HL7 standards. By regulating in this way what would be allowed to be communicated via subsequent v2-style HL7 messaging systems, v2's problems of dialect formation would, it was hoped, be solved.

3. HL7's Act-Centered View of Healthcare

The RIM starts out from the assumption that any profession or business, including healthcare, can be viewed as consisting primarily of a series of *intentional actions* on the part of responsible actors working within an organizational framework. The varieties of such intentional actions relevant for healthcare include:

assessment of health conditions (such as the taking of your pulse, or the weighing of your baby);

provision of treatment services (such as performing surgery, or administering drugs);

informing of patients and their next of kin about health conditions;

provision of notary services (such as the preparation of a living will);

editing and maintaining of documents;

ordering and accepting delivery of supplies;

reporting to government agencies;

billing;

and many more.

Interestingly, HL7 explicitly acknowledges the influence of philosophy in its creation of the RIM, whose act-centered view of healthcare draws for its underlying framework on the speech act theory developed by J. L. Austin in Austin (1962) and by John R. Searle in Searle (1969).

Austin and Searle were among the first philosophers to recognize the theoretical significance of the fact that *what we can do with words* goes well beyond uses of language of the statement-making sort.[1] We can make requests, issue commands, make promises, ask questions, and so on, and actions of these sorts are marked by the fact that the very utterance of words brings about some extra-linguistic result, as for example when the making of a promise brings about the result that the maker of the promise stands under a certain obligation to perform a certain act.

Speech acts of different types, now, can share the same propositional content. Thus I can command that you open the door; I can suggest that you open the door; I can ask whether you will open the door; and so on. Moreover, as was recognized by pioneers of the logic of action,[2] there is a sense in which this same propositional content can be shared also by actions of a non-linguistic sort, as when a command is obeyed or an instruction is followed. In virtue of this sharing of contents speech acts and other human actions form certain standard sorts of sequences, as when, for example, a question is followed by an answer, an act of transmitting information is followed by an act of acknowledgement, a promise is followed by the performance of the promised act, or an act of ordering bedding supplies is followed by acts of, for example, transporting, receiving, billing and paying for the bedding supplies delivered.

The RIM sees this idea of shared propositional content as an attractive way of modeling how the domain of human actions is organized. In sequences such as the sort described, we are to utilize the same RIM 'classCode' — for example 'replenish bedding supplies!' or 'register this patient!' or 'administer this drug!' — to capture the common content of what is involved in each successive act within the sequence, combined with a succession of different 'moodCodes' (such as 'order' or 'command') to capture what is peculiar about each succeeding act. The 'model of healthcare information' that is created in this way is seen by the authors of the RIM as providing an efficient and reliable framework for ensuring successful communication of meaning within and between healthcare information systems.

Of course there are many features of healthcare that go beyond the category of action. These include the *participants* of the actions them-

---

[1] For a broader view of the history of speech act theory see Barry Smith (1990).
[2] See for example G. H. von Wright (1963).

selves, both agent and patient. They include the *roles* these participants play in actions, including their *authority* to perform given actions. They include the sorts of entities to which these actions give rise (Smith 2003), such as *obligations*, *claims* and *electronic documents*. They include *diseases*, and the associated causal processes inside the organism, including processes such as birth and death. They include *material objects* such as pharmaceutical products, DNA samples, equipment and buildings, and they include *organizations and institutional entities* such as insurance companies, government agencies, and laboratories. All of these must be taken into consideration in a complete ontology of the healthcare domain. Ultimately, an ontologically adequate language for communication of healthcare information should have the resources to capture all of the items on this list and all of the different sorts of relations that hold between them, and to do this in a maximally intuitive way that is at the same time easily modifiable as the needs and practices of healthcare organizations change with time.

4. The RIM Straightjacket

HL7's current documentation of the RIM standard appears in an International Standards Organization (ISO) document entitled "Health informatics: — HL7 version 3 — Reference information model — Release 4", and described as a "Draft International Standard (DIS)".[1] The document serves as the basis for a ballot to establish the new Release 4 of the RIM as an ISO standard.[2]

As the document makes clear, the RIM requires that all healthcare information will be organized in terms of just the six "backbone" classes presented in Table 1 below. As the reader will see, this Table comprises two lists, of descriptions, and of definitions, which seem (to us, at least) to be in various ways inconsistent, even though they are taken from a single document. It is difficult to write clearly about HL7 specifications when these specifications are themselves formulated in inconsistent ways. Moreover, there are a number of further problematic issues with the individual entries in these two lists.

---

[1] Health informatics: – HL7 version 3 – Reference information model – Release 4 (Document: ISO/HL7 21731:2011(E)), http://standardsproposals.bsigroup.com/Home/getPDF/1361, last accessed November 28, 2012.

[2] As we understand the matter, the proposal is to establish the RIM as an international standard for being the RIM.

First, they are marked by an embarrassing circularity, as for example in the definitions of 'Entity' and 'Role', which can be of no possible assistance to someone who does not already understand what HL7 takes to be the meaning of the terms defined.

Second, are problems of ambiguity, for example when we are told that *Act* 'represents the actions that are executed and must be documented as health care is managed and provided' does this mean that actions voluntarily recorded do not fall under the heading of Act? Clarity, here, would demand a distinction, to which we shall return below, between health-care actions in general and actions of documentation in particular.[1]

Backbone Class	Description (from 0.2: "RIM as an abstract model")	Definition (from 9.3: "Code System")
Act	represents the actions that are executed and must be documented as health care is managed and provided	a record of something that is being done, has been done, can be done, or is intended or requested to be done.
Participation	expresses the context for an act in terms such as who performed it, for whom it was done, where it was done, etc.	indicates that the target of the participation is involved in some manner in the act, but does not qualify how.
Entity	represents the physical things and beings that are of interest to, and take part in health care	a physical thing, group of physical things or an organization capable of participating in Acts while in a role.
Role	establishes the roles that entities play as they participate in health care acts	a competency of the Entity that plays the Role as identified, defined, guaranteed, or acknowledged by the Entity that scopes the Role.
ActRelationship	represents the binding of one act to another, such as the relationship between an order for an observation and the observation event as it occurs	a directed association between a source Act and a target Act.

[1] Thus although HL7 is uncertain in its habit in this respect, we shall for purposes of clarity normally capitalize the first letter of HL7 terms such as 'Act', 'ActClass', and so forth. When terms such as 'entity', 'act', 'action', and 'role' appear without initial capitals in this essay (other than in quotations from HL7 documents), then their common meanings are intended.

RoleLink	represents relationships between individual roles	a connection between two roles expressing a dependency between those roles and permitting the authorization or nullification of a dependent role based on status changes in its causal or directing role. The RoleLink may be operated over time and thus whose state and identity must be managed [*sic*].

Table 1: The Six Backbone Classes of the HL7 Reference Information Model

Unfortunately the RIM's definition of Act positively undermines a distinction along these lines by identifying 'Act' with 'Record' — and this, even though the definitions of Entity and Role provided in the same ISO document see the latter not as records, but rather as the Entities and Roles themselves, thus further consolidating the 'incoherence' which we identified in the RIM already in 2006 (Smith and Ceusters 2006).[1] Third, are problems of interpretability, as when the document oscillates — to us mysteriously — between the use of 'act' and 'Act', or 'role' and 'Role', sometimes within one and the same sentence.[2]

Fourth, and perhaps most importantly, is the apparent narrowness of scope (in a standard 'reference information model' that is intended to cover the entire domain of healthcare). The RIM's list of backbone classes is intended to be exhaustive, yet important families of items seem be excluded. Above all, where is the place for diseases and for disease processes inside the patient's body? Where is the place for hospital-related adverse events such as falls or spills or leakages of radioactive materials? These are not Acts, they are not contexts for Acts, they are not Entities, and they are not Roles. Where, then, do they fit within the RIM?[3]

In what follows we discuss our attempt to make sense of the RIM's backbone classes in terms of what they *include*, and drawing on HL7's own documentation and usage. We exploit in this connection certain fundamental ontological categories distinguished by philosophers in

---

[1] See Barry Smith and Werner Ceusters (2006: 133–138).

[2] See for example the definition of 'Disciplinary action' as: 'An action taken with respect to a subject Entity by a regulatory or authoritative body with supervisory capacity over that entity. The action is taken in response to behavior by the subject Entity that body finds to be undesirable.'

[3] We discuss HL7's response to this question in Smith and Ceusters (2006), *op. cit.*

dealing with speech acts and similar phenomena, categories which reflect, within a systematic, logical framework, the central common-sensical distinctions such as that between a thing and an event, or be-tween what is particular and what is general. We believe that any frame-work of definitions that cannot be cashed out intelligibly in terms of such distinctions will not be teachable to, and learnable by, normal hu-man beings, and thus will likely lead to errors and confusions (and thus to the very sorts of inconsistent development which were responsible for HL7's problems of dialect formation).[1]

For this purpose we utilize as our instrument of evaluation the Basic Formal Ontology (BFO), an upper-level ontology originally developed in IFOMIS and now used by many groups of researchers throughout the world as a vehicle for promoting interoperability of systems designed for handling scientific and many other sorts of data.[2] We select BFO as framework for our assessment of the RIM because it contains a set of categorizations which have been both well-tested from many different perspectives and also carefully defined and elucidated from a logical point of view.

5. BFO: Independent Continuant

We use 'entity' (with lower case 'e'), in what follows, as an ontological term of art comprehending all items (objects, things, features, attributes, patterns,…) that exist in any way. (HL7's 'Entity' is thus much narrower in its extension.) All real-world entities, from the BFO perspective, for example all entities of the sort that we encounter in the domain of healthcare, fall into one of two exclusive categories of *continuant* and *occurrent* (Grenon, Smith and Goldberg 2005).

Continuants are entities which continue to exist through time; they preserve their identity from one moment to the next even while under-going a variety of different sorts of changes. Continuants are divided by BFO into the two sub-categories of *independent* and *dependent con-tinuant*, the latter being distinguished by the fact that they depend for their existence on the former in the way in which, for example, the tem-

---

[1] "Are the ISO 21090 Data Types Too Complex?", http://hl7-watch.blogspot.com/2010/11/are-iso-21090-data-types-too-complex.html, last accessed November 12, 2012. This is just one example of multiple posts at this site documenting the RIM's unteachability.

[2] http://www.ifomis.org/bfo/, last accessed November 12, 2012.

perature or mass of a material body dependence on this material body for its existence.

Typical examples of independent continuants from the healthcare domain include human beings, buildings, wheelchairs, scalpels, and paper documents in filing cabinets. Each of these entities continues to exist through time even as it undergoes changes, for example, a human being will continue to exist, and preserve its identity, even as it grows and ages over time.

The RIM's Entity seems, at first sight, to be a close analogue to what BFO identifies as independent continuants. Entities are described by the RIM as 'physical things and beings that are of interest to, and take part in health care'. However, when we examine some of the subtypes of Entity in the RIM, as illustrated in Table 2, we find a number of items which are not physical in the normal meaning of the word.

Most blatant is 'Imaging Modality', which (in conformity with standard usage among radiologists) is asserted to be a subclass of 'Device'.[1] It is, however, *defined* by the RIM as: 'Class to contain unique attributes of diagnostic imaging equipment'. This is to confuse a piece of equipment with one or more of the attributes of a piece of equipment. This confusion may well be compatible with the RIM's *description* of Entity (see Table 1), if the latter is to be read as having '(physical entities) and (beings)'[2] as its intended scope and if we are allowed to include attributes as 'beings'. But it is incompatible with the RIM's *definition* of 'Entity', since an attribute of a physical thing is neither a physical thing, nor a group of physical things, nor an organization capable of participating in Acts while in a role.

Worryingly, this example strongly suggests also that, when HL7's authors are formulating definitions, they have no sure understanding of the meaning of the very word 'definition'. In the case of 'Organization' we are indeed provided with a statement of necessary and sufficient conditions and thus with a definition of a logically recognizable sort. In the case of 'Imaging Modality', however, we are provided with something like a statement of the reasons why those responsible for introducing a term thought it necessary to do so (as if one were to define 'screw-

---

[1] http://hl7.org, last accessed November 12, 2012.
[2] As contrasted with 'physical (attributes and beings)'.

Entity =def. A physical thing, group of physical things or an organization capable of participating in Acts while in a role.

> Living Subject =def. Anything that essentially has the property of life, independent of current state (a dead human corpse is still essentially a living subject).

> Health Chart Entity =def. A health chart included to serve as a document receiving entity in the management of medical records.

> Organization =def. A social or legal structure formed by human beings.

> Group =def. A grouping of resources (personnel, material, or places) to be used for scheduling purposes. May be a pool of like-type resources, a team, or combination of personnel, material and places.

> Place =def. A physical place or site with its containing structure. May be natural or man-made. The geographic position of a place may or may not be constant.

> Material = def. Any thing that has extension in space and mass, may be of living or non-living origin.

>> ManufacturedMaterial =def. An Entity or combination of Entities transformed for a particular purpose by a manufacturing process.

>>> Device =def. A subtype of ManufacturedMaterial used in an activity, without being substantially changed through that activity.

>>>> Certificate Representation =def. A physical artifact that stores information about the granting of authorization.

>>>> Imaging Modality =def. Class to contain unique attributes of diagnostic imaging equipment.

Table 2: HL7 RIM Entity class and selected subclasses[1]

driver', for example, as meaning 'a term I will need next week when I reach the letter "S" in my list').

Another problem example in Table 2 is 'Health Chart Entity', an immediate subclass of RIM 'Entity' that is defined as follows:

Health Chart Entity =def. A health chart included to serve as a document receiving entity in the management of medical records.

---

[1] Health informatics: – HL7 version 3 – Reference information model – Release 4 (Document: ISO/HL7 21731:2011(E)), *op. cit*.

Because the definition is circular, there is no easy way to understand its meaning. Is a Health Chart Entity a person or agency that can 'receive' a document? If so, then why is it not included as a child of 'Person or Organization'. Or does 'Health Chart Entity' refer to the document itself? If so, then this would bring the implication that such a document is, in accordance with the RIM's definition of 'Entity', as a 'physical thing, group of physical things or an organization capable of participating in Acts while in a role'. To conceive a document as a physical thing, however, creates problems in view of the fact that the documents of interest to a healthcare organization will in many cases be electronic documents, and thus information artifacts (abstract patterns created through special processes which may be stored simultaneously on many different devices). As we have argued elsewhere, this is an item of a sort which calls for a treatment quite different from that of physical entities, a treatment that must distinguish clearly between the device that stores information and the information entities that are stored.[1] That the RIM does not acknowledge this distinction is seen in its treatment of 'certificate representation', which is defined as 'A physical artifact that stores information about the granting of authorization' and is asserted at the same time to be a subtype of 'device', defined as a subtype of material, and thus as 'having extension in space and mass'. What is the *mass* of an e-certificate granting authorization?

Further problems are raised by the class 'living subject', which is for some reason *not* treated as a child of 'entity class material', even though the latter is defined by the RIM as 'Any thing that has extension in space and mass, may be of living or non-living origin'. The problems here are compounded still further when we are told that 'a dead human corpse is still essentially a living subject'.

6. BFO: Dependent Continuant and the RIM: RoleClass

Dependent continuants, in BFO, are the *states*, *properties*, *qualities*, and *roles* of patients, administrators and so forth. The category of dependent continuant is particularly important for an understanding of the ontology of social reality. Examples of special relevance for us here are the *mental* and *normative* states to which some actions give rise, including

---

[1] http://code.google.com/p/information-artifact-ontology/, last accessed November 12, 2012.

above all the *intentions* of the participants on the one hand and their *obligations* and *claims* on the other. The category of dependent continuants includes also the *capacities* and *skills* (counted as *dispositions* in BFO) of healthcare personnel, such as the ability to speak Spanish or to perform complex medical procedures; the *roles* — for example the nurse or patient roles — that participants play in actions; and their *authority* to perform given actions in virtue of having these roles. These entities are continuants in the sense that, like organisms and molecules, they preserve their identity over time. For example, an intention is a state; that is, it is something that endures from point of inception to point of realization. An entity of this sort is *dependent* in the sense that it requires the support of at least one other entity — its bearer — in order to exist. A *relation of authority* is similarly a dependent continuant, in this case of a sort that has a multiplicity of bearers, namely (i) the human being who has the authority in question, and (ii) the human being(s) over whom this authority is wielded.

Some dependent continuants are captured in the RIM by the classes *Role* and *RoleLink*. In the RIM, an Entity which participates in an Act must do so in a particular Role. The Role defines the Entity's competency (which actions it can perform) and constraints (which actions it cannot perform). In some cases, the Role connects the player of the Role to those bodies, groups, or agencies that have the power to recognize the Role. An example from the RIM is *LicensedEntityRole*, which is a relationship in which, for example, a medical authority certifies a medical caregiver as being permitted to perform certain activities that fall under the jurisdiction of the medical authority in question. The RoleLink class defines connections between Roles. Examples include *has direct authority over* and *has indirect authority over*. In this way the RIM comprehends chains of authority in an organization. While there are some similarities between HL7 Roles and BFO dependent continuants, however, there are also significant differences, to one of which we now turn.

From the BFO point of view, the universals represented in an ontology are instantiated by particulars. Universals correspond to the general terms (such as 'cell' or 'electron' or 'pneumonia') used in scientific texts and also to the even more general terms (such as 'independent continuant' and 'entity') used in high-level, domain-neutral ontologies such as BFO that have been designed to support the data annotation and

cross-domain data integration needs of scientists. Particulars are the entities that we can observe, for example in the lab or clinic, or record, for example when we register a baby's weight or a nurse's promotion.

Particulars then instantiate corresponding universals (Smith and Ceusters 2010). Just as humans, hospitals, kidneys, and so forth, instantiate the universals *person*, *hospital* and *kidney*, so dependent continuants such as Pippa's weight, or John's nurse role, or Mary's authority over Harvey, instantiate universals such as *nurse role* and *authority*, respectively. Just as it is common in a healthcare setting to find multiple instantiations of the universal *human*, so it is common to find multiple instantiations of the same *role*. For example, *nurse practitioner role* is multiply instantiated whenever a hospital has more than one nurse practitioner; but in each case it is the same role universal that is being instantiated. In the case of an authority role, the picture might look like this:

	independent continuant	dependent continuant
universal	*organism*	*role*
	person	*authority role*
particular	Mary, Harvey	Mary's authority over Harvey

Between Mary and the universal *person*, and between Mary's role as someone who has this specific authority in this specific healthcare institution, and the universal *authority* there obtains the relation of *instantiation*. Between Mary's authority and Mary herself there obtains the relation of *inherence* (Mary herself is the *bearer* of her authority). Between *person* and *organism* and between *nurse role* and *role* there obtains the relation of *subsumption* (*person* is a subtype, or sub-universal, of *organism*).

Can we now identify 'Role' in the RIM with what is called 'role' in BFO? Unfortunately not. Again, the extension of Role has been conceived by the authors of the RIM in a seemingly arbitrary way, with the result that it is too ontologically heterogeneous to identify with any class of entities that has been coherently defined. Certainly some subclasses of

HL7 Role represent dependent continuants in BFO's sense, as for example in cases such as this:

Healthcare provider	An entity (player) that is authorized to provide health care services by some authorizing agency (scoper).
Patient	A Role of LivingSubject (player) as a recipient of health care services from a healthcare provider (scoper).
Health cart	The role of material (player) that is the physical health chart belonging to an organization (scoper).

Table 3: Examples of subclasses of HL7 Role[1]

But there are on the other hand also specializations of Role which have nothing to do with roles as commonly conceived. Consider what the RIM calls *RoleClassOntological*, including *has generalization*, *instance* and *subsumed by* as subtypes (see Table 4).[2] The latter, unfortunately, have nothing to do with *roles* in the commonsensical usage of this term.

RoleClassOntological		A relationship in which the scoping Entity defines or specifies what the playing Entity is. Thus, the player's "being" (Greek: ontos) is specified. [Description: RoleClassOntological is an abstract domain that collects roles in which the playing entity is defined or specified by the scoping entity.]
	Subsumed by	Relates a prevailing record of an Entity (scoper) with another record (player) that it subsumes. Examples: Show a correct new Person object (scoper) that subsumes one or more duplicate Person objects that had accidentally been created for the same physical person. Constraints: Both the player and scoper must have the same classCode.
	Has generalization	Relates a specialized material concept (player) to its generalization (scoper).
	Instance	An individual piece of material (player) instantiating a class of material (scoper).

---

[1] Health informatics: – HL7 version 3 – Reference information model – Release 4 (Document: ISO/HL7 21731:2011(E)), *op. cit*.
[2] RoleClassOntological is introduced by the RIM as an immediate subclass of Role, but in a separate group from Patient, Caregiver, and other roles, alongside RoleClassAssociative and RoleClassPartitive.

Equivalent entity	Specifies the player Entity (the equivalent Entity) as an Entity that is considered to be equivalent to a reference Entity (scoper). The equivalence is in principle a symmetric relationship, however, it is expected that the scoper is a reference entity which serves as reference entity for multiple different equivalent entities. Examples: An innovator's medicine formulation is the reference for "generics", i.e., formulations manufactured differently but having been proven to be biologically equivalent to the reference medicine. Another example is a reference ingredient that serves as basis for quantity specifications (basis of strength, e.g., *metoprolol succinate* specified in terms of *metoprolol tartrate*.)
Same	The "same" role asserts an identity between playing and scoping entities, i.e., that they are in fact two records of the same entity instance, and, in the case of discrepancies (e.g. different DOB, gender), that one or both are in error. Usage: playing and scoping entities must have same classCode, but need not have identical attributes or values. Example: a provider registry maintains sets of conflicting demographic data for what is reported to be the same individual.

Table 4: The RIM's RoleClassOntological and examples of its subtypes[1]

The RIM's authors have, it would seem, for some reason come to the conclusion that it is necessary to add basic ontological terms and relations (such as 'kind', 'instance', 'subsumed by', and so forth) into the framework of the RIM. The recommended approach, in such circumstances, would be to adopt the best practices common in the appropriate discipline, which is in this case the discipline of ontology. (And, for all their differences, there is a great deal of commonality among the three most widely used upper-level ontologies — namely BFO, DOLCE and SUMO — as concerns basic ontological terms and relations.) Unfortunately, as so often been in the history of HL7, the RIM's authors chose to develop a special, idiosyncratic framework of their own, even though this meant sacrificing interoperability with externally developed standards. Moreover, they chose to do this on the basis of the assumption that the terms and relations in question must be subsumed under the RIM's existing six backbone classes.

---

[1] Health informatics: – HL7 version 3 – Reference information model – Release 4 (Document: ISO/HL7 21731:2011(E)), *op. cit.*

The result, we are sorry to say, is an egregious potage of confusion. To view *subsumed by* as a role is analogous to viewing the relation between, say, the species *rabbit* and the genus *mammal* by conceiving the latter as a role played by the former. And worse: all of the RIM's backbone classes, and indeed the entire structure of the RIM, require for their understanding the prior distinction between particular and universal. It does not make sense to conceive this distinction itself — under the heading 'Instance' — in terms of Roles.

7. BFO: Occurrent and the RIM ActClass

Occurrents (also called events or processes) are defined from the BFO point of view as being items which are such that they unfold themselves in their successive phases. Thus in contrast to continuants, occurrents never exist in full in any single instant of time.[1] The life of a patient is an occurrent, as also is the course or history of a given disease or of a given treatment. Actions are occurrents, and so also are sequences of actions, from planning, to issuing of orders, to the execution of a plan. (Plans themselves however are continuants: thus they endure continuously through time until they reach the point of complete execution or abandonment.)

Act represents the closest analogue in the RIM to occurrents, but it is at best a weak analogue, since as we have argued elsewhere (see Smith and Ceusters 2006), it is defined both too narrowly and too broadly, to a degree that gives rise to the hypothesis that (as in the case of 'Role') it is not capable of being coherently defined. Here we pursue the assumption that — in keeping with the conception of the RIM as a 'model of healthcare *information*', rather than of the *reality* which such information is about — Act does not comprehend actions themselves, but rather only the records which arise when actions are documented by a healthcare professional in either a clinical or an administrative context. These records themselves are (for the RIM) Acts. (That such an identification brings confusion to the user is revealed most poignantly in the fact that the RIM / ISO document referred to above itself contains multiple passages, including the definition of Entity quoted in our Table 1, in

---

[1] We leave aside here the case of occurrent boundaries, for example beginnings and endings.

which 'Act' is used to mean, not 'the record of an act', but rather the act itself, in which some Entity participates.)

To help us to understand the rationale behind the RIM's design, it is useful to point to the view of the medical record on which it is based, a view formulated by Rector and Nolan in 1991, according to which the medical record is a record, not of *what is the case on the side of the patient,* but rather of 'what clinicians have said about what they have heard, seen, thought and done' (Rector, Nolan and Kay 1991). From this point of view, not faithfulness to the clinical history and care of the patient is the fundamental criterion for what gets included in record, but rather, as Rector and Nolan express it:

> The first consequence of our view of faithfulness is that the information in the medical record itself is not about what was "true" of the patient but what was observed and believed by clinicians.[1]

Consider the case where physician A documents (at time *t* and place *p*) that physician B obtains a blood sample from patient C. Whether or not physician A actually documented that 'physician B obtains a blood sample from patient C' is of vital importance to the medical record. What is of lesser importance, according to the Rector-Nolan view, is whether or not the proposition 'physician B obtains a blood sample from patient C' is true. The idea is that information about the real world can be brought into the medical record only through records — *descriptions of acts* — formulated by suitably authorized persons. Since these descriptions are always attributed to someone, it is possible to have a medical record which contains statements about one and the same real world activity which yet disagree in their propositional content (for example because their authors were provided with conflicting information). The medical record will then still be consistent, for it will contain descriptions not of the form

physician B obtained a blood sample from patient C

physician B did not obtain a blood sample from patient C

but rather of the form

---

[1] On the significance of the use of "snear" quotes around the word 'true', here, see Stove (2011).

authorized person A_1 recorded: 'physician B obtained a blood sample from patient C'

authorized person A_2 recorded: 'physician B did not obtain a blood sample from patient C'.

This means, however, that the extension of the RIM's Act — meaning the totality of acts in reality described through Acts — overlaps only partially with that of the BFO category of occurrents. This is because Act, for the RIM comprehends not only the records which result from documenting processes *that have happened*, but also counterpart records of processes that *did not happen*, that *can happen*, *are intended to happen*, *are requested to happen*, and so forth, in reflection of the RIM's distinguished mood codes. Or as HL7 itself formulates the matter: 'The moodCode distinguishes among Acts that are meant as factual records, records of intended or ordered services, and other modalities in which acts can be recorded.'[1] Acts will thus include also, for instance, (records of) intended but cancelled surgeries. The fact that there can be Acts, in the sense of the RIM, which never happen is from this point of view understandable. Unfortunately however this outcome is in conflict with the RIM's own description of Act (see Table 1, again) as representing *the actions that are executed*.

8. Conclusion

Our purpose has been to show how the RIM's backbone classes line up, or rather fail to line up, with the fundamental categories of the BFO ontology. If BFO, as we believe, comes close to capturing categorical distinctions at the heart of common sense, then this failure implies major shortcomings in the RIM — shortcomings for example as concerns learnability, coherence, comprehensiveness, and stable evolution. As we saw, the HL7 has presumed that the problems of dialect formation which had plagued v2 would be resolved once the RIM was brought into play. On the basis of this presumption the HL7 organization has gone on to promulgate manifold varieties of RIM-conformant 'v3' standards, and these new standard — even in spite of their manifest shortcomings — have been incorporated into multiple health information technology

---

[1] Health informatics: – HL7 version 3 – Reference information model – Release 4 (Document: ISO/HL7 21731:2011(E)), http://standardsproposals.bsigroup.com/Home/getPDF/1361, last accessed November 28, 2012.

initiatives, where they have sometimes been associated with conspicuous failures, some of them on a national scale.[1] On the other side, however, there are also positive signs which are slowly but surely beginning to manifest themselves, and we are pleased to observe that the claims made on behalf of the RIM are increasingly being treated with suspicion in healthcare informatics circles.[2]

References

Austin, J. L. (1962). *How to do Things with Words*, Oxford: Oxford University Press.

Grenon, P., Smith, B., and Goldberg, L. (2005). "Biodynamic Ontology: Applying BFO in the Biomedical Domain", in D. M. Pisanelli (ed.), *Ontologies in Medicine*, Amsterdam: IOS Press, 20–38.

Grieve, G. (2012). "HL7 needs a fresh look because V3 has failed", http://www.healthintersections.com.au/?p=476, last accessed November 12, 2012.

Rector, A., Nolan, W., and Kay, S. (1991). "Foundations for an electronic medical record", *Methods of Information in Medicine*, 30:179–86.

Searle, J. (1969). *Speech Acts. An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.

Smith, B. (1990). "Towards a History of Speech Act Theory", in A. Burkhardt (ed.), *Speech Acts, Meanings and Intentions. Critical Approaches to the Philosophy of John R. Searle*, Berlin/New York: de Gruyter, 29–61.

Smith, B. (2003). "John Searle: From Speech Acts to Social Reality", in B. Smith (ed.), *John Searle,* Cambridge: Cambridge University Press, 1–33.

Smith, B. and Ceusters, W. (2006). "HL7 RIM: An Incoherent Standard", *Studies in Health Technology and Informatics*, 124:133–138.

---

[1] http://hl7-watch.blogspot.com/2011/08/uk-scrapping-national-health-it-network.html, last accessed November 12, 2012.

[2] Grahame Grieve, "HL7 needs a fresh look because V3 has failed", http://www.healthintersections.com.au/?p=476, last accessed November 12, 2012.

Smith, B. and Ceusters, W. (2010). "Ontological Realism as a Methodology for Coordinated Evolution of Scientific Ontologies", *Applied Ontology*, 5, 139–188.

Stove, D. C. (2011). *Scientific Irrationalism: Origins of a Postmodern Cult*, New Brunswick: Transaction.

von Wright, G. H. (1963). *Norm and Action: A Logical Inquiry*, London: Routledge & Kegan Paul.

Intentionality and Indexicality: Content Internalism and Husserl's *Logical Investigations*

Andrew D. Spear

1. Introduction[1]

A traditional approach to issues of meaning, reference, experience, and knowledge is to analyze them in terms of the intentionality of consciousness. On this view intentional events consist of three distinct but correlated components: the intentional act, the intentional content, and the intentional object. On such a view, the object that an intentional event is about is determined by the intentional act and the intentional content of that event. Further, on this view intentional events are characteristically conception-dependent, object-independent, and content-indeterminate (Smith & McIntyre 1982). 'Conception-dependence' means that thought about an object is always thought of the object as being a certain way or by means of a specific description (e.g. Napoleon *as* the victor at Jena, a triangle *as* a three-sided geometrical figure). 'Object-independence' means that it is possible to have meaningful thoughts about objects that do not exist (e.g. Pegasus, phlogiston). And 'content-indeterminacy' means that the way in which a subject thinks about an object rarely, if ever, involves a complete description or determination of all features of the object (e.g. *S*'s thinking about the evening star does not guarantee that the content of *S*'s thought also includes that the evening star is the morning star or that the evening star is the planet Venus; *S* may in fact not be thinking of or even know these things at all). In addition to garnering support from basic phenomenological observations, the traditional view of intentionality is also supported and recommended by the fact that it can explain a number of traditional puzzles from the philosophy of language and mind, including informative identity statements,

---

[1] I first met Ingvar Johansson in Buffalo, NY, but truly became acquainted with him during a year at the Institute for Formal Ontology and Medical Information Science at the University of the Saarland in Germany in 2006. Working as colleagues, and regularly talking over lunch at the Imbiss Café, I was struck by Ingvar's wide philosophical knowledge, his seemingly inexhaustible enthusiasm for philosophy, and his generosity with his time and comments. I have learned a great deal from Ingvar and am honored to contribute to this Festschrift in honor of his seventieth birthday.

meaningful thought about non-existent objects, and failure of inter-substitution in intensional contexts.

This view of intentionality is a kind of descriptivism and a kind of internalism about mental content insofar as it is intentional content that determines what object a thought is about by presenting that object as being a certain way. This understanding of intentionality is thus committed to some version of the thesis that intentional content is the "sense" of significant thoughts and assertions and that it is this intentional content that determines the referent or extension of the thought, a characteristic internalist thesis. It is also plausible to view intentional content as determined by internal non-relational features of intentional subjects. If content determines reference and for the subject to have a certain content is just for her to be in a certain mental state, then it seems right to say that the content is determined by the mental states of the subject only. Further, on this view it is possible for a subject to have an intentional thought whose object does not exist, even though the thought has content. In such a case all that could determine the content of the thought are internal features of the subject.

As a kind of content internalism, this view of intentionality is subject to the challenges that have been raised in the last half decade by various types of content externalism and accounts of direct reference. The primary arguments against content internalism itself, which have focused on natural kind terms, indexicals, and demonstratives, proposes cases where it seems clear that two subjects are qualitatively identical regarding the psychological states they are in and the intentional contents of their thoughts but who, contrary to the principle that content determines reference, are clearly thinking about or referring to *different* extensions or objects. Such cases are supposed to force the internalist to choose between the thesis that content determines reference and the thesis that content is determined by internal psychological states of the subject. The thesis that content determines reference is a central one for descriptivist internalism insofar as this just is the explanation of how thought is directed toward objects or extensions. However, rejecting the thesis that content is determined by internal features of the subject opens up the possibility that external features of the object of thought or of the subject's environment play a role in determining what the content of a subject's thought is, even when the accessible descriptive content involved

in that thought remains constant. But accepting the possibility that the content of thought might include non-descriptive external elements undermines the explanatory power of the content-object distinction in addressing traditional puzzles of meaning and reference. Thus, an internalist committed to the traditional conception of intentionality has reasons to resist giving up either of these commitments.

In what follows, I will consider the Twin Earth thought experiment of Hilary Putnam concerning natural kind terms and similar cases that have been proposed by John Perry for indexicals. I will argue that these cases function as definitive arguments against content internalism, understood in terms of the traditional account of intentionality, only on the assumptions that (i) internalism is committed to what I will call a strict Fregean interpretation of the content determines reference thesis and (ii) that there is no descriptive content associated with demonstrative and indexical thoughts (or at least not enough to establish their reference). I will argue that Edmund Husserl's *Logical Investigations* understanding of the ontology of intentional content provides a framework within which it is possible to defend a content internalism committed to a weaker content determines reference principle, and that identifies intentional content associated with indexical and demonstrative thoughts, while retaining the explanatory advantages of the content-object distinction. Such a view is immune to standard externalist objections and retains the explanatory power of traditional theories of intentionality.

2. The Traditional Conception of Intentionality

The "traditional conception of intentionality" is a way of thinking about mental states, their contents and their objects to be found in the work of thinkers such as Franz Brentano and Edmund Husserl, as well is in the writings of more recent philosophers such as Roderick Chisholm, John Searle and Tim Crane (Brentano 1995; Husserl 2000; Crane 2001; Chisholm 1981; Searle 1983). The basic idea of intentionality is that thought consists of a correlation between mental acts and the objects that they are about: it is of the essence of thought and of significant experience more generally to be about or directed toward an object. While the initial statement of the view is simple enough, its further development leads to a number of insights, each of which is a point of departure for further theorizing about the structure of thought and experience. Three features have played a central role in almost all accounts of inten-

tionality in this tradition: these are that intentionality is (i) "existence-independent", (ii) "conception-dependent", and (iii) that it admits of indefiniteness in what it represents.[1]

Intentional thoughts are existence-independent because they can be about objects that do not exist, either in cases of error or in cases where the thought simply is about something non-existent (such as phlogiston, the present king of France, or a round square). Intentional relations are conception-dependent insofar as the mind is, at least phenomenologically speaking, never simply related to an object *per se*. Rather it is always related to an object from a certain perspective or under a certain description or way of thinking about that object. In perceiving, the object is always perceived from a certain perspective and under a certain conception or description. In thinking, an object is thought of under a certain conception or in a certain way; one can think of Napoleon as "the victor at Jena" or as "the vanquished at Waterloo", but it does not seem that one can think of Napoleon while not thinking of him as being *any* way at all. Finally, intentional presentations can involve indeterminacy insofar as they need not, in their content, fully specify all features of the object they are about. Thus, the thought that "John is a doctor" is about John with respect to the question of whether or not he is a doctor, while leaving open, simply not being about or making a decision one way or another with regard to many other features of John, such as his height, whether he is married, what kind of person he is, etc. This indeterminateness is often a feature of intentional thought about things such as John, but *not*, of course, a feature of John himself.

2.1 Existence-Independence

Viewing intentionality as existence-independent is motivated by the "paradox of intentionality". If intentionality is a normal relation between a thought and an object, and the existence of such a relation entails the existence of its relata, and some intentional objects do not exist, then it seems to follow that *some non-existent objects are objects* (Crane 2001: 23). To spell this out more explicitly,

(i)   Intentionality is a relationship between a thought and an object.

---

[1] This terminology is taken from Smith & McIntyre (1982: chapter 2), however the same distinctions are to be found, using different terminology, in Searle (1983) and Crane (2001: chapter 1) among others.

(ii) If a relationship exists, then its relata also exist.

(iii) Some intentional thoughts are about objects that do not exist.

Since by (i) intentionality involves a *relationship* between a thought and an object, then by (ii) any time an intentional relation exists both a thought and an object (its relata) must exist as well. But, by (iii) some intentional thoughts are related to objects that do not exist.

C) Therefore some objects are objects that do not exist.

While some philosophers have been willing to accept this conclusion and to try to incorporate non-existent objects into their ontologies as the real relata of intentional relations,[1] most have seen the conclusion of this argument as paradoxical and have argued that one of the premises needs to be rejected or altered. Since it is obvious that some thoughts are about things that do not exist and it seems equally clear that the existence of a relation (such as "giving birth to", "striking" or "causing") *does* entail the existence of its relata, the most straightforward way of resolving the paradox is to modify the first assumption in some way; the assumption that intentionality is always a relationship between thought and an object. Once this is done it is possible to either view intentional states "adverbially" (when John thinks about Pegasus he is "thinking Pegasusly" just as when John experiences green he is "seeing greenly", etc.) or to maintain that intentional states are relations to something other than their objects. For example, Frege views intentional relations as obtaining between a thought and an abstract sense, thus for him the obtaining of this relationship does not require the existence of a corresponding object in all cases (Frege 1948). What these moves have in common is that they involve making a distinction between the content (or sense or adverbial mode) and the object of an intentional thought, so affirming

---

[1] The most notorious example being Meinong (1960). For a general discussion of non-existent objects, see It is worth noting that, whereas Meinong was willing to accept the conclusion of this argument, Frege modified premise one, effectively replacing 'object' with *Sinn* or sense, while Russell in "On Denoting" retains premises one and two and avoids the Meinongian conclusion by denying three, effectively maintaining that there are no meaningful thoughts about non-existent objects (thoughts that seem to be about such objects, e.g. "the present King of France", are really about something else, namely the complex relations amongst quantified groups of objects and properties referred to by definite descriptions) (Frege 1948; Russell 1905).

the view that intentional states are independent of the existence of the objects they are about.

2.2 Conception-Dependence

The conception-dependence of intentionality is supported by a number of factors. First, phenomenological observation reveals that perception is perspectival. When one intends an object perceptually one always experiences it from a particular perspective or vantage point. Thought about objects seems the same way: one does not simply think about an object, one thinks about it *as* some kind of object or *as* determined in a certain way or *in relation to* some other object or objects. Perceptions and thoughts don't just present the world, they present it *as being a certain way.*

In addition to phenomenological data, there are the traditional puzzles of informative identity statements ("Hesperus is Phosphorus") and of the failure of intersubstitutability *salva veritate* in intensional contexts (e.g., Lois believes that Superman, but not Clark Kent, can fly). The fact that a subject can know something about an individual described in one way, but fail to know about the same individual when the individual is described in another way, strongly suggests that thought about objects is conception-dependent in the way under consideration here: thought about an object is always thought about an object *as* something or *in a particular respect.*

2.3 Indefiniteness and Indeterminateness

Finally, the indefiniteness or indeterminateness of intentionality is the flipside of the conception-dependence of intentionality. If every thought about an object is a thought about that object under a certain mode of presentation, then any given thought about an object will *determine* some things about that object, present it as being a certain way, while leaving others open and undetermined. What is determined is definite and rules out some other ways of thinking about the object, but what is indeterminate leaves open and even positively suggests certain possibilities of future determination of the object in thought or experience.

Some thoughts will be *indefinite* in the sense that they do not even specify a particular object that they are about (for example, the desire for a glass of water is, usually, not a desire for any particular glass of water, and for most of us, a thought about "the world's tallest woman" is not a

thought about any particular person), but even thoughts that are about a specific object under a specific conception will still be *indeterminate* with regard to a large number of the object's qualities. (E.g. to think of Napoleon as "the victor at Jena" is to determine that he is a human being, a military leader of some sort and, with respect to this particular battle, the victor; but thinking this thought does not, by itself, determine anything about the question of Napoleon's other qualities, such as his appearance or height, or how he faired in other military endeavors, such as Waterloo.) Thus, on the traditional conception of intentionality, the complete identity of the object of thought is *more* than or *more* finely individuated than the content of any given thought about that object *presents it as* on a given occasion.

2.4 The Act-Content-Object Structure of Intentionality

In light of the foregoing, intentional events can be analyzed in terms of four components: a cognitive subject, the act of intending itself, the object that is intended, and the content or *way* in which the object that is intended is thought about by the subject. Thus, every intentional event involves a *subject* presenting to herself or directing herself toward an *object* in a certain *kind of way* and by means of a *conception* or way of thinking about that object.[1] If Joan perceives the top of a box, then Joan is the *subject* of the intention, perceiving is the particular kind of *act* of intending she is engaging in, a "box viewed-from-the-top" is the *content* or way in which she perceives and recognizes the object, and the *object* of her intention itself is the entire three dimensional physical object that is the box. This yields the following schema for intentional events, intentional: Subject → Act → Content → Object.[2] I will be assuming this structure and using this terminology in what follows.

---

[1] The use of 'intention' here has no special connection with the idea of an agent 'intending' to do something or with intentional action. As it is used here, 'intention' indicates, roughly, what many philosophers today refer to as propositional attitudes.

[2] Searle uses "psychological mode" for intentional act, and "representational content" for intentional content (Searle 1983), while Tim Crane speaks of "intentional mode" and "intentional content" (Crane 2001) and, in the *Logical Investigations* (Husserl 2000) Husserl used "act-quality" for intentional act and "act-matter" for intentional content, while later in Ideas (Husserl 1982) he used "noesis" and "noema" for these two notions respectively. For discussion of this terminology and its

For my purposes here the subject of intentional events will not receive much attention. What will be more important is that there are systematic *correlations* between intentional acts, intentional contents and intentional objects: it is possible for different *acts* to have the same content (it is possible to perceive a red apple, to wish for one, to remember one, etc.), and for different *contents* to be directed toward or about the same object (as in the case of the box used as an example above, it is possible to perceive it first from one side, then from another; it is also possible to think of, say, Napoleon *as* "the victor at Jena" or *as* "the vanquished at Waterloo", or the number 2 first *as* "even" and then *as* "the successor of 1").

2.5 The Content-Object Distinction: Contents as Quasi-Descriptive Senses

The distinction between the content and the object of thought is of fundamental importance for the theory of intentionality. The content of a thought is the specific *way* in which a given intentional event presents some object as being and plays the role of directing the mind of the subject towards that object. On the view under consideration here, intentional contents are understood as quasi-descriptive senses,[1] similar in function to the Fregean notion of *Sinn* (Frege 1948). This way of viewing intentional content is arguably also part of Husserl's conception of intentionality, and has been endorsed by David Woodruff Smith and Ronald McIntyre, as well as by John Searle in his notion of the "conditions of satisfaction" for intentional content (Smith and McIntyre 1982: 206; Searle 1983). To say that intentional contents are quasi-descriptive senses is to say that they have something like truth-conditions. Intentional contents *present* the world as being a certain way, and it is thus in virtue of their contents that they are about or directed to what they are about while it will be in virtue of the way the world *is* that such thoughts will succeed in referring to an object, or in being true or false with regard to one. In other words, thought contents mean or "say" that the world is a certain way, while the objects and states of affairs in

---

development in Husserl, see Smith and McIntyre (1982: chapter 3) and Simons (1995).

[1] I say "quasi-descriptive sense" here because, as I will explain below, the sense involved in intentional content as I am using it here need not be or involve explicitly linguistic expressions or content.

the world determine whether or not what these contents "say" or mean is indeed satisfied by or accurate of the world.

3. The Traditional Conception of Intentionality as Descriptivist and so Internalist about Mental Content

The traditional conception of intentionality is, arguably, a version of content internalism, though it is difficult to pronounce on the issue definitively, since this way of thinking about the mind historically predates discussions of internalism and externalism about mental content. In "The Meaning of 'Meaning'", Hilary Putnam characterizes the traditional understanding of meaning as committed to two theses:

> (I)   That knowing the meaning of a term is just a matter of being in a certain psychological state.
>
> (II) That the meaning of a term (in the sense of "intension") determines its extension (in the sense that sameness of intension entails sameness of extension). (Putnam 1975:136)

For Putnam, the significance of (I) is ultimately that it implies that "…the psychological state of the speaker determines the intension (and hence, by assumption (II), the extension)" of a term (1975: 139). Now, if 'meaning' and 'intension' are replaced with 'content', as many philosophers have felt free to do, then the traditional account of content is, according to Putnam, committed to the thesis that (I) the content of a thought is determined only by the psychological states of the subject of that thought and (II) the content of a thought determines the referent or extension of that thought, what it is about. The view of intentionality under discussion here is plausibly a version of the traditional account of content as Putnam understands this.

The traditional conception of intentionality is most clearly committed to some version of descriptivism, the view that a subject succeeds in thinking about or referring to something in virtue of possessing descriptive information of some sort that is adequate to pick out the object or extension apart from all others.[1] Intentional content is a kind of descript-

---

[1] As Husserl writes, "The matter [intentional content] must be *that element in an act which first gives it reference to an object, and reference so wholly definite that it not merely fixes the object meant in a general way, but also the precise way in which it is meant*…It is the act's matter that makes its object count as this object and no other…" (italics in the original)(Husserl 2000:589).

ive content, even if the descriptions or representations that it involves will not always be linguistic. There are a number of motivations for this descriptivism. First, positing intentional contents consisting of descriptive information is a way of beginning to answer the basic question of why some things in nature are intentional, have thoughts and beliefs that are about other things, while others are not. Merely causal and physical relations are pervasive in nature, both among beings that are and beings that are not intentional, so positing a distinctive component of thought, its involving descriptively structured intentional content, begins to explain the difference between intentional and non-intentional beings. Second, as already noted, there is phenomenological evidence for descriptivism. When objects are perceived or thought about, this perceiving and thinking is in fact accompanied by some descriptive information, some mode of presentation or other. Third, it seems that in order for a subject to be thinking about an object, she must have some understanding of what it is that she is thinking about. It seems implausible to say that S is thinking about O, but has no conception whatsoever of what O is in virtue of which she picks it out from other possible objects of thought. Suggesting that S thinks of objects in virtue of possessing identifying descriptive information about them avoids this implausibility. Fourth, and also noted above, distinguishing between descriptively structured intentional content on the one hand and the object of thought on the other makes it possible to provide unified responses to traditional puzzles of meaning and reference, such as the puzzle of informative identity statements. As a version of descriptivism, the traditional conception of intentionality is thus committed to some version of Putnam's (II), which I will characterize here as:

> *Content Determines Reference* (CDR): The content of a thought determines the referent or extension of that thought, what it is about.

What about Putnam's first thesis, that the content of a thought is determined by the psychological states of the subject of that thought only? To say that the content of a thought is determined by psychological states of the subject is to say that where content is internal, two subjects who are internal mental duplicates, having all historical and current mental properties and experiences in common, will also be content-duplicates, they will be instantiating qualitatively identical mental contents.

The idea that content is entirely determined by the psychological states of cognitive subjects does seem to be a commitment of the traditional conception of intentionality, albeit a less direct one than CDR. Accepting it is motivated both by the existence-independence of intentionality and by its conception-dependence. If it is possible for a subject to have a thought (act and content) where the object of that thought fails to exist, then it seems clear that that thought *must* supervene on the intrinsic features of the subject; what else could it depend on for its existence? The conception-dependence of intentionality also motivates commitment to content internalism insofar as it suggests that the mode or way in which a subject is directed towards the object of thought depends on that subject. Further, it seems to be the case (as in illusions or hallucinations of perception) that two intentional acts could have qualitatively identical content, even though the object of one of these acts fails to exist. If this is correct, then it seems that the contents of such thoughts depend on the psychological states of the subject alone (since in the hallucinatory case there is nothing else for the content to depend on). Thus, the traditional conception of intentionality seems committed to:

> *Psychological States Determine Content* (PSdC): The content of a subject's thought is determined by the psychological states of the subject only.

The traditional conception of intentionality, being committed to CDR and PSdC, is clearly committed to versions of the theses about meaning that Putnam ascribes to traditional views of meaning, and so will also be susceptible to Twin Earth and other thought experiments meant to show the incompatibility of these theses, and so the falsity of this traditional internalist view. In the following section I will consider some of the most prominent examples of these challenges and identify certain common assumptions about content internalism that they make.

4. Externalism as a Challenge to the Traditional Conception of Intentionality

Recently, it has been argued by many that the conjunction of the two commitments of content internalism just discussed, the content determines reference principle and the thesis that psychological states determine content, have been decisively refuted, and that this motivates com-

mitment to externalist views of content according to which the content of a thought is determined by at least some factors other than the psychological states of the subject whose thought it is. In what follows my goal is to show how the traditional view of intentionality is, in fact, able to accommodate the challenges posed by Twin Earth and other anti-internalist arguments.

The basic structure of anti-internalist arguments is to propose situations in which two subjects are internally identical and so identical with respect to the contents of their thoughts, but where it seems intuitively clear that the reference or aboutness of the subjects' thoughts is distinct, thus they are using the *same* internalist content to refer to or think about *different* objects. If this is right, then either the content determines reference principle or the psychological states determine content principle must be abandoned. What is most important about these externalist scenarios, what gives them their force against internalism, is that the subjects described in these scenarios seem to be clearly indistinguishable with regard to the descriptive content, broadly interpreted, that is available to them. What the scenarios seem to show is that the descriptive contents that subjects have available are simply *not sufficient* to explain why the subject's thoughts are about what they are about.

Now, while such cases are indeed a challenge to traditional internalism, I will argue that what they really draw attention to is the way in which features of *context* can play a role in determining what a subject is thinking about at a given time. For example, what the thought "I am here now" refers to will depend on who speaks it, where, and when. The internalist is committed to saying that there is a quasi-descriptive content grasped by subjects who think this thought, but must at the same time accommodate the fact that the referents of the thought will be different on different occasions. To do this, the internalist must do two things. First, articulate a version of the content determines reference principle that leaves open the possibility that type-identical mental content tokens may, in different contexts, refer to or be about different objects. Second, provide an explanation of the descriptive content or "conditions of satisfaction" involved in such contents that makes context-sensitivity possible (Searle 1983: chapter 8).

In what follows I will focus on Hilary Putnam's Twin Earth argument and John Perry's discussion of cases involving indexicals and show how

each of these anti-internalist arguments rests on a certain reading of the content determines reference principle and on a seeming inability of content internalism to make sense of indexical and demonstrative contents. However, it is my position that most if not all of the other anti-internalist arguments admit of analyses along the same lines as the ones that I will provide here.[1]

4.1 The Externalist Challenge and the Need for an Internalist Account of Indexicality

The basic argument of Hilary Putnam's Twin Earth thought experiment is that if (i) content determines reference and (ii) content is determined by psychological states of the subject only, then two psychologically identical subjects thinking the same thought in different contexts should be referring to *exactly* the same thing or extension regardless of the difference in context. Since it is possible to generate cases of psychologically identical subjects thinking the (descriptively) same thought in different contexts where, due to features of the context itself, they seem clearly to be referring to different objects or extensions (H_2O and XYZ respectively in Putnam's thought experiment), one of the two principles must, Putnam argues, be rejected. Here is a more thorough formulation of Putnam's Twin Earth argument:

---

[1] The exception to this, it could be argued, is Burge's "social anti-individualism", the locus classicus for which is Burge (1979). Burge's argument that the content of individuals' thoughts involving terms such as 'arthritis' can be externally individuated by the meaning for such terms that is accepted in their linguistic community (such that two individuals using the same term with the same internal 'content' nevertheless mean or refer to different objects or extensions in virtue of being situated in different linguistic communities) is more complicated than standard anti-internalist arguments insofar as it relies on a discussion of concept possession and partial concept possession. My own position is that how one interprets Burge's arguments for social anti-individualism will depend more or less directly on whether one takes the more basic kinds of Twin Earth style arguments themselves to be conclusive. If they are, then Burge's argument for the social case is a natural extension, while if they are not (as for internalist sympathizers such as myself), then something like Searle's account of "parasitic intentionality" will be the natural response, see Searle (1983:250). At any rate, I will not be pursuing this kind of anti-internalist argument at any length here. Thus I note the omission.

(i) *Content Determines Reference:* Assume that (a) content determines reference to at most one object or extension and (b) content is determined by the subject's psychological states.

(ii) *Possibility of Internal Duplicates:* It is possible for there to be internal mental and physical duplicates. Individuals who are qualitatively identical with respect to all of their intrinsic properties, including their psychological states.

(iii) *Logic:* If (i) and (ii) then any two internal duplicates will be thinking about/referring to exactly the same objects at any time at which they have identical psychological states.

(iv) *The Twin Earth Intuition*: But Earth individual and Twin Earth individual are internally identical agents who, by hypothesis, have internally identical thoughts and are yet thinking about/referring to *different* objects (the extensions containing H_2O and XYZ respectively).

C) Therefore, by *modus tollens*, (i) is false and either (a) or (b) must be rejected.

Putnam takes the balance of his argument to show that it is best to reject (b), the principle that psychological states determine content, though he admits that doing this renders the import of the content determines reference principle "vacuous" as well (Putnam 1975:165). The upshot is that external features, usually of the subject's environment, must play some role in determining content and, ultimately, in establishing reference or aboutness as well.

Now, Putnam's argument relies essentially on the way in which subjects can use demonstrative thought and pointing in a context to secure reference to something, even when they know very little about the nature of that thing. This can be seen by considering how one standard internalist line of response to Twin Earth fails to succeed. This line of response is to simply maintain that, so long as the two subjects are really internally identical, the extensions of their thoughts includes *both* XYZ and H_2O. After all, for all these subjects know prior to the discovery of the microstructure of water, the comprehensive descriptions that they associate with 'water' are equally true of both H_2O and XYZ, even if they have never come into contact with one of these two substances.

What makes this line of response difficult for the internalist to maintain is Putnam's emphasis on the idea that 'water' and other natural kind

terms are generally introduced ostensively, by pointing to the surface features of a particular stuff in one's environment and committing to thought and reference about the yet undiscovered microstructure of *that very stuff*, whatever it is. As Putnam writes,

> My "ostensive definition' of water has the following empirical pre-supposition that the body of liquid I am pointing to bears a certain sameness relation (say, x is the same liquid as y, or x is the same as y) to most of the stuff I and other speakers in my linguistic community have on other occasions called "water". (Putnam 1975: 141)

This way of viewing how natural kind terms get their meaning makes the internalist line of response just mentioned much more difficult to maintain. For any natural kind term that gets its meaning in *this* way (and it can't be denied that there simply aren't *any*), it will be possible to propose a mental duplicate in a qualitatively identical environment whose act of ostensive definition for that kind term clearly picks out something with a different microstructure. Hence the Twin Earth objection to internalism will survive any internalist attempt to argue that content determines reference to both Earth and Twin Earth water because it can always be insisted that some, perhaps even most, kind terms get their meanings in this ostensive way. Putnam himself is very clear about this 'indexical' component in natural kind terms (Putnam 1975: 153). While natural kind terms do not function *just* like indexicals, insofar as it seems unlikely that their extensions remain sensitive to context once their reference has been initially fixed, it remains the case that the difference in the extension of the contents associated with 'water' for Earth and Twin Earth subjects is to be explained by the fact that the reference of this term has been fixed demonstratively to paradigm instances of water that differ dramatically in underlying microstructure.

What the Twin Earth argument shows is thus that demonstrative reference in a context can make a difference to what an individual succeeds in referring to or fixing reference to. So long as it is possible to provide a cogent internalist account of the demonstrative reference fixing of meaning for natural kind terms, there is no reason why an internalist could not appeal to these differences in the fixing of the reference of the terms to explain the difference in content between Earth and Twin Earth subjects. When Earth subjects use 'water', the content they associate with it is that it is anything relevantly similar in underlying microstructure to the stuff

with respect to which the reference of the term was originally fixed. Since this stuff will be different for Earth and Twin Earth subjects, it should be no surprise if their later intentions to refer to "anything relevantly similar to *that* very stuff", while internally the same, yet refer to different extensions, even if these subjects do not know this fact (internalism is definitely *not* the thesis that subjects are omniscient about the objects that their thoughts refer to). Everything depends on the demonstrative.

4.2 John Perry on Indexicals

A similar point can, not surprisingly, be made with regard to two of John Perry's arguments involving indexicals. The first argument is similar in structure to Putnam's Twin Earth argument. Perry proposes Hume and his Twin Earth doppelganger Heimson, who both believe the proposition "I am David Hume" (Perry 1977:487–90). If (as the traditional Fregean view seems to hold) sense determines reference regardless of context and understanding a sense amounts to being in a psychological state, then since the two individuals are doppelgangers and so type-identical, they will presumably express the same proposition when they respectively utter tokens of the sentence 'I am David Hume'. But, argues Perry, the truth-values of what is said in the two cases, as well of course as what the expression 'I' refers to, are clearly different. Since the descriptive contents of Hume and Heimson's thoughts are the same, but the referents are different, it must be something more than internal states or descriptive content that determines reference in these cases.

Similar to the Twin Earth argument, what Perry's Hume-Heimson case shows is that indexical thought in a context can make a difference to what an individual succeeds in referring to, even if most or all of that individual's internal mental content is descriptively the same. The challenge for the internalist is to provide a plausible internalist account of indexical thought content, one that explains how indexical thoughts have conditions of satisfaction that establish reference to *different things in different contexts*.

Perry's second argument is for the essentiality of indexical statements (Perry 1979). Perry maintains that, whatever the meaning of indexicals such as 'I' are, sentences containing them cannot simply be paraphrased by sentences containing purely descriptive non-indexical terms. For example, 'I am writing a letter' cannot be paraphrased as 'John Smith is

writing a letter on May 5, 2008 and etc.' without loss or alteration of some content. To show this, Perry relies on the reasonable premise that belief plays a role in determining action, and then develops scenarios in which belief in an indexical sentence plays a role in motivating an individual to take action, whereas belief by that same individual in a descriptive non-indexical paraphrase of the same sentence *does not* play the same role in motivating the individual to take action *unless* she also possesses the original indexical belief. To illustrate this, Perry tells a story about himself at the grocery store. He is following a trail of sugar through the isles, trying to find the person who has a broken bag of sugar in their cart so he can let them know. Eventually, he looks down and realizes that the broken bag of sugar is in his own cart, which leads him to adjust it so that it will stop making a mess. The moral to be drawn from this story, according to Perry, is that the belief that "someone is making a mess" and even the belief that "John Perry is making a mess" would not be sufficient by themselves to motivate an individual to take action *unless* that individual also believed that "I am the person making a mess" or that "I am John Perry". This suggests that there is a dimension of *content* for indexicals such as 'I' that cannot be entirely captured by sentences involving only non-indexical terms. Why is this a problem for content internalism?

The content internalist is committed to a term→content→object model. Terms and sentences have their significance in virtue of being associated with an intentional content, which in turn determines the object or extension they are about. Thus a word, such as 'water' expresses a meaning or sense "the wet stuff, H_2O, etc.", which determines reference to an extension, construed as all of the particular stuff of which the descriptive elements of the sense are true. However, this view in conjunction with the content determines reference principle raises a problem for indexicals. For if an indexical such as 'I' does indeed express a sense, the question is "what is it"? One traditional answer has been to say that the meaning or sense of 'I' is "the very person who is speaking". Thus, indexicals get glossed as expressing non-indexical third person descriptions and sentences containing indexicals are taken to tacitly express neutral third-person propositions where the indexicals are replaced with proper names (for 'I', 'you'), specific dates (for 'now',

'yesterday'), locations (for 'here', 'there') and etc.[1] Sentences containing indexicals thus have a different meaning, express a *different* content, on each occasion of their use, and so preserve the spirit of the strong sense determines reference principle. But it is precisely this move that Perry's argument for the essentiality of indexicals blocks.

Perry's argument shows that there is a component of sense for indexical expressions that is *lost* when the indexicals in such expressions are paraphrased away in favor of third-person descriptive content. Now, once again, this does indeed raise a problem for an internalist committed to the content determines reference principle, for such an internalist *must* try to explain how the senses or contents involved on different occasions of indexical use are indeed *different* insofar as each succeeds in referring to a different thing, but this difference cannot, it seems, be a difference in the descriptive content associated with different occasions of use.

4.2 Singular Thoughts and Demonstrative Reference

Finally, it is possible to offer relatively straightforward anti-internalist arguments involving demonstratives. Summarizing arguments of this kind, Jessica Brown (2004:13–15) proposes a case of two internal duplicates each of whom is having the qualitatively identical experience of an apple before her, but where the apples are non-identical in the two cases. In a case such as this, the perceptual contents (and so the contents of any demonstrative reference based on them) seem to be clearly identical, but the referents in the two cases are just as obviously different. The content determines reference principle is once again challenged, so internalism seems to fail for demonstrative contents as well.

4.4 Content Determines Reference & Indexical and Demonstrative Content

A major supposition of anti-internalist arguments is that the content determines reference principle is a strict one:

---

[1] An example of this sort of response to the problems posed by indexicals can be found in the final chapter of Cohen and Nagel (1993).

Strict Content Determines Reference (SCDR): content determines reference and each content determines at most one referent or extension.[1]

This supposition is fair, historically speaking, as Frege himself was clearly committed to it. According to Frege, the meanings of linguistic terms are senses, which are abstract particulars. It is the sense or thought expressed by a linguistic expression that determines what, if anything, it refers to, and since senses are abstract particulars, each sense is only able to establish reference to (at most) one object or extension. This model works well for many kinds of terms, especially the more abstract terms of mathematics, logic, and the sciences, and can also be applied to proper names. However, this account runs into trouble with indexicals and demonstratives.

The reason is that the content of such expressions seems to have two parts. On the one hand, 'I', 'you', 'here', and 'this' surely have some common core of meaning or rule of use that is constant from one usage to another. The word 'I' does not have a completely different meaning for John and Dan when each of them utters 'I am hungry', even if it does have a different referent. But the common meaning associated with indexicals is not by itself sufficient to establish what they refer to on a given occasion. If the common sense or meaning of indexicals was all that was involved in establishing their reference, then by the principle that meaning determines reference, they should always refer to the *same* thing. But, of course, they don't. So it seems that there must be some additional "completing sense", different for each occasion of use of an indexical, which makes it possible for the subject's thought to succeed in referring to, e.g., her current location apart from all others. Further, this "completing sense" cannot consist merely of third person descriptions of the subject's current location, for due to the points raised by Perry about the essential nature of indexicals, thinking that "it is hot in the capital of Italy" is not the same as thinking "it is hot here" unless one is also aware that "here is the capital of Italy", which reintroduces the indexical and so the problem of providing a "completing" indexical sense for the subject to grasp. But now it begins to look like, for each occasion of the use of

---

[1] 'At most' because it is possible, as already noted, to have contents without referents. E.g. phlogiston.

an indexical (or at least for each person, place, time, etc.) there must (i) be a distinct completing sense that establishes the referent of the indexical on that occasion, but (ii) these completing senses cannot be composed entirely of neutral third person descriptions, for if they were then they would still not, in light of Perry's discussion of the essentiality of indexicals, capture the full meaning of indexical expressions, so these senses must involve an indescribable (in any words other than the appropriate indexicals) component that can only be thought by the appropriate subject or at the appropriate time or in the appropriate place for each distinct indexical completing sense. But, understood in this way, commitment to the existence of a very large number of partially inexpressible abstract indexical senses seems implausible.

To accommodate the case of indexicals (and demonstratives) what and the internalist needs to do is provide a different view of the metaphysics of intentional content (one that does not appeal to abstract Fregean senses) that allows for the articulation of a version of the content determines reference principle that leaves open the possibility that type-identical mental content tokens may, in different contexts, refer to or be about different objects. Doing this will make it possible to provide an explanation of the descriptive content or "conditions of satisfaction" involved in indexical and demonstrative contents that is the same on different occasions of use, but that nevertheless due to its context sensitivity and its being actually instantiated in a particular context, refers to different objects on different occasions. Significantly, the view of the metaphysics of intentional content and the partial account of indexicals offered by Edmund Husserl in the Logical Investigations holds out the promise of doing both of these things.

5. Husserl's LI View of Intentionality as a Version of the Traditional Conception that Can Survive the Challenges of Externalism

In the *Logical Investigations* Husserl developed a view according to which conscious acts are primarily intentional, and a mental act is intentional just in case it has an act-quality, an act-matter and an act-character. The quality of an act is the kind of act that it is, whether perceiving, imagining, judging, wishing, etc. The matter of an act is what I have been calling its content, it is the mode or way in which an object is thought about, e.g. a house intended from one perspective rather than another, or Napoleon thought of first as "the victor at Jena", then as "the

vanquished at Waterloo". The character of an act includes such things as whether it is an act of merely reflecting on a possibility (a "non-positing act") or one of judging or asserting that something is the case (a "positing act"), as well as the degree of evidence that is available to support the intention of the act as fulfilled or unfulfilled (as genuinely presenting some object in just the way that the act-matter suggests, or not). The notion of act-character is important for purposes of epistemology, but here I will abstract from it in order to focus, primarily, on Husserl's notion of content or act-matter.[1]

Husserl's notion of content as act-matter is different from the standard Fregean notion discussed above. Whereas the standard Fregean view sees the sense of a thought as an abstract particular that the thought must somehow *grasp*, Husserl views act-quality, act-matter and act-character as mutually dependent constituents of a concrete particular thought itself. Just as there cannot be color without saturation, brightness and hue, so for Husserl there cannot be an intentional act without quality, matter and character. The act-matter or content of an act, according to the early Husserl, is a real dependent part of the intentional act itself rather than an abstract particular of some sort that the act must reach out and grasp. Thus, on Husserl's view it is less mysterious how a subject has access to the sense or content of her intentional state insofar as the act-matter is a literal constituent of that state itself.

Whereas Fregean accounts deal with the fact that one individual can have the same thought at different times and different individuals can think about the same thing at any time by positing a *single* abstract sense that is the numerically identical content of all of their thoughts, Husserl views particular act-matters or contents as instances of ideal *act-content species*. Thus, on Husserl's view, two subjects are able to think about the same thing in the same way when both of them instantiate *exactly similar* instances of a single kind of content or act-matter. Thus if John

---

[1] Husserl's own development of this view occurs in Husserl (2002: Investigations I [for discussion of meaning and reference], V [for the structure of intentionality itself], & VI [for the discussion of the epistemological implications of the previously developed views]). One of the most sustained discussions of the metaphysics of meanings as intentional contents that is my focus here is in *Logical Investigation* I, Chapter Four (Husserl 2000: 328–333). A helpful overview of Husserl's views is Simons (1995).

and Sarah are both thinking about how they would like to see the Twins win the 2008 World Series in baseball, they are having the same thought and thinking about the same objects in virtue of instantiating exactly similar act-matters of the single act-matter type "the Twins win the 2008 World series in baseball" (the hoping that this comes about would fall under act-quality rather than act-matter). On this view of intentional content, it is not the abstract content species or types, but rather specific instantiations of them in the thoughts of intentional subjects that determine reference. While this fact may make little difference for abstract thoughts in mathematics, logic, and well-developed areas of science, it will make a great deal of difference for cases where the intentional content itself might involve context sensitive indexical or demonstrative elements. This suggests a modification of the content determines reference principle as follows:

> *Logical Investigations Content Determines Reference* (LICDR): The instantiation of a content-type in a thought determines the referent or extension of *that* thought, what it is about.

If it is possible to provide an account of the descriptive information or "conditions of satisfaction" for indexical and demonstrative contents such that different instantiations of the same "content-species" would refer, as required, to different objects in different contexts of use, then the traditional theory of intentionality understood as committed to something like LICDR will not be affected by Twin Earth style anti-internalist arguments. In light of the arguments considered so far, especially those of John Perry, is such an account possible?

5.1 A Basic Internalist Account of Indexical Thought

It is well recognized by those who wish to defend internalism about content in the context of an account of the intentionality of the mind that (i) the main issue is that of providing an adequate account of indexical and demonstrative thought, and that (ii) doing this involves modifying the strict principle that sense always *uniquely* determines reference along the lines of LICDR above (Smith and McIntyre 1982: chapter IV, section 3; Smith 1984; Searle 1983: chapter 8; Crane 2001: chapter 4). Appealing to Husserl's *Logical Investigations* understanding of intentional content has already made clear how it is possible to think of intentional content as establishing reference in a context, insofar as on this view it is only

instantiations of intentional content types that actually succeed in referring. The question that remains and that must be answered in order to use this view of the content determines reference principle to deal with cases of indexical and demonstrative content is: what is the descriptive or description-like content of essentially indexical thoughts?

Here too the views of the early Husserl point toward an answer. In the *Logical Investigations* Husserl recognized the need for a distinction between what he called "objective" expressions on the one hand, and those that are "essentially occasional" on the other."[1] According to Husserl, essentially occasional expressions include both indexicals and demonstratives, and such expressions have *two* facets of meaning. The first is what Husserl calls a constant "semantic function" associated with particular indexical expressions. For example, "It is the universal semantic function of the word 'I' to designate whoever is speaking…" (Husserl 2000:316). John Searle calls such semantic functions "lexical meanings" and identifies them with a non-indexical descriptive component associated with indexical expressions (Searle 1983:224–5). Searle maintains that the lexical meaning of an indexical expression will specify both the *kind* of relationship that must obtain between a speaker's utterance and features of the context of her utterance (spatial proximity, relations of proximity in time, conversational direction, etc.), as while as what *kinds* of things are to be related to the utterance (instants of time, the speaker, the listener, locations, etc.). Thus 'you' lexically means "the very person (kind of thing) being addressed by (kind of relation) this utterance". Husserl recognizes, as does Searle, that for the reasons suggested by Perry's discussion of the essential indexical, the sentences expressing these semantic functions or lexical meanings cannot simply be substituted for indexicals without affecting the meaning of sentences containing them (Perry 2000:315). This makes it necessary to identify a second facet or component of indexical content.

---

[1] According to Husserl, an objective expression is one that, "…pins down (or can pin down) its meaning merely by its manifest, auditory pattern, and can be understood without necessarily directing one's attention to the person uttering it, or to the circumstances of the utterance" (Husserl 2000:314). An essentially occasional expression by contrast is, "…an expression that "belongs to a conceptually unified group of possible meanings, in whose case it is essential to orient actual meaning to the occasion, the speaker and the situation" (Husserl 2000:315).

To deal with this, Husserl proposes a distinction between the semantic function or "indicating meaning" of indexicals, which remains constant from use to use, and the "indicated" meaning of indexicals, which is fundamentally cued to certain features of the speaker and context of utterance. Thus the "indicating meaning" of 'I' is always "whoever is now speaking", but the *indicated* meaning of its use on a given occasion is keyed to the "self-awareness" or "self-presentation" of the speaker on that occasion. In general, the indicating meaning of an indexical will specify some general relationship between the utterance of a sentence and some feature of the speaker's conscious awareness or perceptually given environment, while the indicated meaning will be determined by what the speaker is *actually aware of* in the context in which the sentence is uttered. In the case of many indexicals, such as 'you' and 'here' their indicating meaning may be supplied in part by demonstrative pointing to features of the immediate perceptual environment. Thus, Husserl writes, "The meaning of 'here' is in part universal and conceptual [semantic function/indicating meaning], inasmuch as it always names a place as such, but to this universal element the direct place-presentation [indicating meaning] attaches, varying from case to case" (Husserl 2000:317–18). John Searle incorporates this same feature in his own account of indexicals by requiring that, in many cases, an "awareness of the context of utterance" on the part of speakers and hearers of indexical expressions will be necessary in order to fully establish their reference (Searle 1983:225–7).

So, here we have a two-part account of indexical thought and reference. Every indexical expression has a general semantic function or lexical meaning which specifies that an utterance of it must stand in a certain relation to a particular kind of thing, such as the speaker, the listener, a time or a place and etc. Since thinking that "the speaker of this sentence is tall" is not, following Perry, the same as believing that "I am tall", an "indicating meaning" or immediate awareness of presentations of features of the context of thought and utterance is identified as a second meaning component necessary in order to specify the full content and hence the reference of an indexical thought on a given occasion. In order for such an account to be an *internalist* account of indexical content, however, more needs to be said about what is involved in the *content* of the awareness of features of the context of utterance. Husserl's

discussion of the "self-presentation" of the speaker as playing the role of "indicating meaning" in a context is helpful, but not sufficiently detailed as it stands. He points in the direction of a solution to the problem, however, when he writes,

> Properly speaking, we should not suppose that the immediate presentation of the speaker sums up the entire meaning of the word 'I'. The word is certainly not to be regarded as an equivocal expression, with meanings to be identified with all possible proper names of persons. Undoubtedly the idea of self-reference, as well as an implied pointing to the individual idea of the speaker, also belong, *after a certain fashion*, to the word's meaning." (Husserl 2000:316)

Husserl does not develop further the insights expressed in this passage, however I think that it is the notions of "self-reference" and of an "implied pointing to the individual idea [content-token] of the speaker" that are crucial in understanding the nature of internalist indexical content. John Searle makes this the central feature of his own internalist account of both indexical and demonstrative content. According to Searle, the contents of indexical thoughts are "self-referential" in the sense that such thoughts *are themselves included* in or made reference to by their own conditions of satisfaction.[1] What this means is that in order for such thoughts to be accurate, in order for the world to be the way which they present it as being, something must be true of that very thought itself. The thought content "this thought is false" is self-referential in this sense. The world will be the way such a thought presents it as being only if something is the case regarding this very thought, namely that it is false.[2] In the case of indexicals, what must be true of the thought itself is that it must indeed stand in the *relationships* specified by the lexical meaning of the indexical thought that it is ('I', 'here', 'now', etc.) to the *kinds of object* the lexical meaning specifies.

---

[1] John Searle's discussion of the self-referentiality of indexical content is remarkably short, and he discusses the matter in terms of speech-acts and utterances rather than in terms of the content of such acts and utterances, which makes it even more difficult to follow exactly what his view of the matter is. The following is an attempt, following Searle, to articulate the matter specifically in terms of content. For Searle's own discussion, see Searle (1983:222–4).

[2] I leave aside here consideration of the paradox that such a thought generates, as it does not apply to the case of self-referential indexical content.

That the appropriate relationship *does* obtain between a thought and the object it is about on a given occasion, however, is something that a subject must determine by reference to her immediate *conscious awareness* of the relevant features of the context of her own thought as well as to her own awareness of the thought itself.

Thus, when a subject thinks the thought "I am hungry" on a given occasion, the subject presents the world as being such that "the thinker of this very thought is hungry" and she understands that she *is* the thinker of this thought based on her immediate acquaintance with or awareness of the thought as hers, something that is indeed irreducibly indexical, but nevertheless internal to the subject's experience.

David Woodruff Smith has proposed a more elaborate account of indexical content along these same lines. Smith challenges the often-made assumption that intentional content of the sort under discussion here must necessarily be descriptive content. Rather, he suggesets that there is another type of intentional content, that involved in experiences of direct acquaintance or intuition such as perception and introspection, and that this content, though it does present the world as being a certain way, does not do so by describing objects as having certain properties. Smith treats indexical expressions as a "*generic* form of acquaintance shared [by thoughts] on different occasions of uttering the term" (Smith 1981: 106). A "form of acquaintance" always involves a kind or structure of intentional experience and reference to whatever kind of object plays the "appropriate" role in that kind of experience, where the appropriate role can be understood as determined by the lexical rule for indexicals of that type (e.g. "the thinker of this very thought in experiences involving 'I', something the subject has access to based on immediate awareness of her possession of her own thoughts; "the time at which this thought is occurring" in experiences involving 'now', which is something that a subject has access to based on her conscious awareness of her location in the flow of subjective time). Where the object that plays the "appropriate" role will be given by what Searle called the "lexical role" for the indexical, though again, actually picking out or establishing reference to this object on a given occasion will require the subject to take account of structural features of her immediate first-person conscious awareness; features that, not surprisingly, will not always be describable in third-person terms.

Thus, maintaining an internalist perspective about intentional content in the case of indexicals requires (i) adopting the LICDR version of the content determines reference principle, (ii) acknowledging the self-referential nature of indexical thought contents, and (iii) recognizing indexical thought contents as establishing their reference, in a given context, based on a more basic but still intentional kind of direct acquaintance that subjects have with their own thoughts, experiences and perceptions. Once this is done, it becomes possible to view indexical thought contents as instantiating a single meaning scheme M_x for each type of indexical thought ("I_x" for 'I', "H_x" for 'here', and etc.). This meaning scheme consists of a general lexical rule, including a self-referential component stipulating the relationship that a thought or utterance of this indexical type must stand in to its object in order for reference to be successful, along with a generic structure or form of conscious acquaintance (such as self-awareness, introspection or perception) that non-descriptively (non-linguistically) presents the world as containing an object or objects that stand in the appropriate relations or play the appropriate role specified by the lexical rule.

On this view, what a token of an indexical type of content (such as 'I') refers to on a given occasion will depend on *whose* thought the combination of the lexical meaning and first-personal acquaintance for that content on that occasion are instantiated in. Thus, when Hume thinks "I am Hume" the self-referential nature of the content involved in his thought in conjunction with his own direct acquaintance with himself refers to him, its conditions of satisfaction are "the person thinking this thought, "I", is Hume", while when Heimson thinks "I am Hume" the self-referential nature of the content involved in his thought is the same but, given that it is instantiated *in his thought content* rather than in Hume's, it *refers to him*. And since Hume and Heimson are indeed different, what Heimson says is false while what Hume says is true, even though they have expressed tokens of a type identical content (Searle 1983:226–7). In other words, if the intentional content scheme for the indexical 'I' is I_x, then both Hume and Heimson think the thought type "I_x am David Hume". However, the content scheme "I_x" only refers once it is embedded in the context of a given thinker's thought and expression and tied to the immediate acquaintance that that thinker has with the structure of his own experience. Since Hume only has immediate access

in the way required by the indexical scheme I_x to his own experience, while Heimson has access only to his own conscious experience, the intentional basis upon which the reference of 'I' is fixed in the two uses is different (one is Hume's and the other is Heimson's) and thus the referents are different as well, even though the content tokens are type-identical, that is to say, qualitatively the same.

The crucial feature of this account of indexical thought content is its appeal to features of immediate experience or acquaintance to ground the reference of self-referential indexical thought in a given context. Though the relevant notion of acquaintance as a non-descriptive form of intentionality requires more development than it has been given here,[1] only its complete unworkability would represent a fundamental obstacle to an account something like the one sketched here. I take it that the guiding idea for Husserl, Searle, and Smith is that each of us has a unique first personal awareness of our own occurrent thoughts and of the way in which they present the world as being. In the case of indexical thoughts, the indexical presentation of the world takes advantage of this first-person awareness of the thought-content as "my" (the speaker/thinker's) content in order to present the world, time, location, and/or other speakers as standing in a certain relationship to the speaker and her current thought. This point about first-personal access is quite arguably supported rather than undermined by Perry's discussion of the essentiality of indexicals, and thus represents, at the very least, a plausible alternative direction to that of taking indexical thoughts to be externally individuated or exhausted by their referents.

Further, since the options available for linguistically expressing indexical content are either to use indexicals, or to eliminate them in favor of non-indexical statements with the same truth-conditions, and since the latter are, due to the unique self-referential and first-personal nature of indexical content, not equivalent in meaning (even if they are equivalent in truth) to the former, it follows that the *full* content of indexical thoughts can only be expressed in sentences containing indexicals, which is fully consistent with Perry's arguments for the essentiality of indexicals.

---

[1] Smith himself has further developed the relevant notion of acquaintance in Smith (1989).

To summarize, an internalist account of indexical content requires rejecting the strict Fregean content determines reference principle. Once this is done, it is possible to understand the meaning of indexicals as given by a meaning scheme "M_x", consisting of their lexical content (a general meaning or semantic function that is always the same and includes the specification of ways in which thoughts with this indexical content must be related to other objects in order to refer) and by the generic form of certain kinds of immediate acquaintance or awareness that subjects can have with features of their experience and environment. Viewing indexicals in this way makes it possible to construe indexical intentional content as a kind of content that is as essentially indexical as the first person nature of acquaintance with one's own experiences is essentially first-personal, a result consistent with and to some extent supported by Perry's discussion of the essential nature of indexicals.

A full account of indexical content, however, such as the content associated with 'here' or 'you', requires that subjects be able to pick out, based on perceptual awareness, objects and features in their surrounding environments. This picking out will be demonstrative in nature, and so an internalist account of indexical content is only complete if there is also some account of internal demonstrative content. It is to the discussion of such an account that I now turn.

5.2 A Basic Internalist Account of Demonstrative Content

As John Perry concludes for Frege, so David Woodruff Smith and Ronald McIntyre conclude of Husserl that he never fully appreciated the problem posed by demonstrative thought and reference, and therefore also never adequately solved it (Perry 1977; Smith and McIntyre 1982). Demonstratives present a problem for an internalist committed to the strict sense determines reference principle because they seem to involve very little internal or descriptive content, and because Twin Earth style cases can be constructed in which use of a demonstrative by internal mental duplicates in different contexts clearly establishes reference to different objects or extensions. As with indexicals, so with demonstrateives, the strategy for providing an internalist account of them is (a) to reject the strict Fregean content determines reference principle in favor of a weaker version such as LICDR, and (b) to develop an account of the nature of internal demonstrative content that ties it to features of the immediate experience of subjects of thought.

Both D. W. Smith and John Searle have offered accounts of what the internal intentional content or sense of perceptions, and by extension of demonstratives, is (Smith and McIntyre 1982:219, ftnt. 32; Smith 1984; Searle 1983:chapters 2 & 8). Smith's basic proposal is that the content of a perception is "the content "this (now here before me and affecting my eyes)" (Smith and McIntyre1983: 219, ftnt. 32). Similarly, John Searle's view is that part of the conditions of satisfaction for a perceptual intention are "the very object that is now causing my perception". Searle refers to this as the causal self-referentiality of the content of perceptual intentions, which is similar to the self-referentiality of indexical contents discussed above. The basic idea in both accounts is that when a subject undergoes a perceptual experience in which she directs her attention at a specific object, part of the meaning of that experience for the subject is that the very object that she seems to see (with whatever qualities she intends it descriptively as having) is *actually* there before her *just as she seems to see it* and *playing a role in affecting her senses and causing* her experience of it. The point is not that most subjects (if any, other than philosophers) actually think *these words* when they undergo perceptual experiences, rather, the point is that it is part of the content of the experience of perception itself, part of the mode or way in which subjects direct their thoughts perceptually towards the world, and so part of the conditions of satisfaction for such thoughts, that the objects so intended are experienced as caught up in and affecting the very perception of them. Since a demonstrative such as 'this' or 'that' is used on the basis of perception (and often in conjunction with some kind of pointing) to establish reference, the meaning of demonstrative expressions can be viewed as depending or relying on the more basic kind of content involved in perceptual intentionality (e.g. "that red round apple" has the conditions of satisfaction, "the red round apple here before me and affecting my experience of it).

On such a view of the intentional content of perception it is not perceptual sense or content alone, but content in conjunction with context that determines completely which object a perceptual intention refers to.[1]

---

[1] The basic picture under discussion here is similar to that suggested by John Heil in his own intuitive response to the implications of Twin Earth for the internalism externalism debate. "Pretend for a moment that the directedness of your thoughts resembled dart tossing. Gravitational influences aside, the direction a dart takes

As Searle puts it, this means that type-identical tokens of the same perceptual state (tokens including all descriptive and qualitative content in common) may yet, in virtue of the causal self-referentiality of perceptual content, still refer to different objects in different contexts (for example, XYZ on Twin Earth as opposed to H_2O on Earth) (Searle1983: 207–9). The same is true on Smith's account, and in virtue of this both the accounts of Searle and of Smith and McIntyre involve accepting the Husserlian content determines reference principle, LICDR.[1] The internalist account of indexical content is thus extended and supplemented by an internalist account of demonstrative content.

5.3 Back to Twin Earth

And this makes it possible to explain what is happening in Putnam's Twin Earth case in a way that is consistent with the standard commitments of internalism, that content determines reference and that psychological features of the subject alone determine content. The reference of a natural kind term such as 'water', understood as Putnam understands it, is determined by (i) demonstrative pointing to a paradigm (ii)

---

depends wholly on agent-centered factors: how you grip the dart, the character of your arm motion, the timing of the release, and the like. Although a dart's trajectory depends wholly on the agent, what the dart hits depends on features of the world, features over which an agent might have no control. When you toss a dart aimed at the center of a target, it will not hit the center if I move the target while the dart is in flight. We might sum this up by saying that what a dart hits depends on two factors: how it is tossed — its agent-determined trajectory — and how the world is" (Heil 2004: 235).

[1] Smith and McIntyre discuss the implications of their view for the strong content determines reference principle in the following passage: "Now, the object of perception is not a function of the content alone, for another perception on another occasion could in principle have the very same phenomenological content and yet have a different object. That is, there is not a functional, or many-one, relation between the content and the object of a perception (contra Husserl). Still, it seems, the demonstrative content of a perception — the content "this (now here before me and affecting my eyes)" — does prescribe the object of the perception, the object appropriately before the perceiver and affecting his sense on the occasion of the perception. However, it is not the noematic content in itself that so prescribes the object; rather, it is the content only insofar as it is embodied in that particular perceptual experience on that occasion — if you will, the demonstrative content-in-the-perception prescribes, or is satisfied by, the object of the perception, the object contextually before the perceiver" (Smith and McIntyre 1982: 219, ftnt. 32).

on the basis of certain superficial identifying features of the paradigm that make it possible to identify other instance of the kind, but (iii) with the intention to refer only to things relevantly similar in underlying microstructure (or some other scientifically significant feature) to the original paradigm. A subject on Earth and a subject on Twin Earth can both instantiate type identical tokens of this content in fixing the reference of their terms 'water', however, the self-referential nature of the demonstrative element of the content ensures that the Earth individual is, in fact, establishing reference to H_2O, while the Twin Earth individual is establishing reference to XYZ. And now all of this is explained by the internal intentional contents of these two subjects.

6. Conclusion

In the foregoing I have introduced the traditional account of intentionality as a kind of content internalism and considered the way in which it is challenged by now standard anti-internalist arguments such as those based on the Twin Earth thought experiment of Hilary Putnam. I have argued that anti-internalist arguments are successful only on a very strict reading of the content determines reference principle, and on the assumption that it is not possible to provide a cogent account of the descriptive or quasi-discriptive content of indexical and demonstrative thoughts. The view of intentional content defended by Husserl in the *Logical Investigations* motivates modifying the content determines reference principle to allow for the possibility that different instantiations of the same content-type may refer to different objects in different contexts, while the conjunction of this view with the development of Husserl's own discussion of indexical and demonstrative thought and reference by John Searle and D. W. Smith, among others, makes it possible to provide a plausible and motivated account of internalist intentional content for indexicals and demonstratives. Content internalism, so understood, is not susceptible to standard anti-internalist arguments and retains all of the explanatory benefits of the traditional distinction between the content and the object of thought as well.

References

Brentano, F. C., Kraus, O., et al. (1995). *Psychology from an Empirical Standpoint,* New York: Routledge.

Brown, J. (2004). *Anti-Individualism and Knowledge,* Cambridge, MA: MIT Press.

Burge, T. (1979). "Individualism and the Mental", in *Midwest Studies in Philosophy IV*, French, Uehling, and Wettstein (eds.), Minneapolis: University of Minnesota Press.

Chisholm, R. (1981). *The First Person: An Essay on Reference and Intentionality*, Minneapolis: University of Minnesota Press.

Crane, T. (2001). *Elements of Mind*, Oxford: Oxford University Press.

Cohen, R. & Nagel, E. (1993). *An Introduction to Logic*, 2nd ed., with introduction by John Corcoran, Indianapolis: Hackett.

Frege, G. (1948). "Sense and Reference", *The Philosophical Review*, 57: 207–30. (Originally published in German in 1892).

Heil, J. (2004). *Philosophy of Mind: A Contemporary Introduction*, 2nd ed., New York: Routledge.

Husserl, E. (2000). *The Logical Investigations*, 2 vols, translated and edited by J. N. Findlay, Amherst, NY: Humanity Books. (1st German ed. in 1900–01.)

Husserl, E. (1982). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*, translated and edited by Fred Kirsten, The Hague: Kluwer. (1st German edition in 1913.)

Johansson, I. (1986). "Levels of Intension and Theories of Reference", *Theoria*, 52: 1–15.

Johansson, I. (2010). "Fictions and the Spatiotemporal World — in the Light of Ingarden", *Polish Journal of Philosophy*, 4: 81–103.

Loeffler, R. (2001). "Demonstrative Reference and Cognitive Significance", *Synthese*, 128: 229–244.

Meinong, A. (1960). "On the Theory of Objects", in R. M. Chisholm (ed.), *Realism and the Background of Phenomenology*, Glencoe, Ill: Free Press. (First published in German in 1904.)

Perry, J. (1979). "The Problem of the Essential Indexical", *Noûs*, 13: 3–21.

Perry, J. (1977). "Frege on Demonstratives", *Philosophical Review*, 86: 474–497.

Putnam, H. (1975). "The Meaning of 'Meaning'", *Minnesota Studies in the Philosophy of Science*, 7:131–193.

Russell, B. (1905). "On Denoting", *Mind*, 14:479–493.

Searl, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.

Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.

Simons, P. (1995). "Meaning and Language", in B. Smith & D. W. Smith (eds.), *The Cambridge Companion to Husserl*, Cambridge: Cambridge University Press.

Smith, D. W. (1989). *The Circle of Acquaintance: Perception, Consciousness, and Empathy*, Dordrecth: Kluwer.

Smith, D. W. (1984). "Content and Context of Perception", *Synthese*, 61:61–87.

Smith, D. W. (1981). "Indexical Sense and Reference", *Synthese*, 49:101–127.

Smith, D. W. and R. McIntyre (1982). *Husserl and Intentionality: A Study of Mind, Meaning, and Language*, Dordrecht: D. Reidel Publ. Co.

Wikforss, Å. (2007). "Semantic Externalism and Psychological Externalism", *Philosophy Compass*, 3:158–181.