



HAL
open science

Singular Terms, Identity, and the Creation of Fictional Characters

Matthieu Fontaine

► **To cite this version:**

Matthieu Fontaine. Singular Terms, Identity, and the Creation of Fictional Characters. *Disputatio*, 2019, 11 (54), pp.207-229. 10.2478/disp-2019-0017 . hal-04390368

HAL Id: hal-04390368

<https://hal.science/hal-04390368>

Submitted on 12 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Singular Terms, Identity, and the Creation of Fictional Characters

Matthieu Fontaine

Center for Philosophy of Sciences of the University of Lisbon

DOI: 10.2478/disp-2019-0017

BIBLID [0873-626X (2020) 54; pp.207–229]

Abstract

How to interpret singular terms in fiction? In this paper, we address this semantic question from the perspective of the Artifactual Theory of Fiction (ATF). According to the ATF, fictional characters exist as abstract artifacts created by their author, and preserved through the existence of copies of an original work and a competent readership. We pretend that a well-suited semantics for the ATF can be defined with respect to a modal framework by means of Hintikka's world lines semantics. The question of the interpretation of proper names is asked in relation to two inference rules, problematic when applied in intensional contexts: the Substitution of Identicals and Existential Generalization. The former fails because identity is contingent. The latter because proper names are not necessarily linked to well-identified individuals. This motivates a non-rigid interpretation of proper names in fiction, although cross-fictional reference (e.g. to real entities) is made possible by the interpretative efforts of the reader.

Keywords

Artifactualism, fiction, world lines, substitution of identicals, existential generalization

0 Introduction

The Artifactual Theory of Fiction (ATF), or creationism, is a realist theory of fictions; that is, a theory which admits fictional characters in its ontology. When talking about ATF, we mainly refer to the theory initiated by Thomasson (1999), who acknowledges fictional characters a full-fledged ontological status, defined by means of the phenomenological notion of ontological dependency. Fictional characters are existing abstract artifacts dependent on various kinds of

entities. On the one hand, Holmes is historically dependent on the creative act of Conan Doyle. Since nobody else could have created Holmes, this dependence is also rigid. On the other hand, Holmes is an abstract entity, dependent on the existence of copies and a competent readership. Since the existence of Holmes is not located in one or the other copy, or competent reader, this dependence is only generic. It is also constant, because if every copy was destroyed, or if nobody was able to read it, or even to remember it, Holmes would cease to exist. The ATF allows a direct account of reference and quantification over fictional entities. Although it is not devoid of metaphysical difficulties, in particular with respect to the creation of fictional characters¹ or the notion of ontological dependency,² we will focus on the semantic question of the interpretation of proper names in literary fictions.

According to Thomasson (1999: 52 ff.), fictional characters are not odd entities. They can be referred to as well as concrete entities. By relying on Kripke's theory of proper names, she considers that a baptism can occur in fiction, and that a chain of reference can travel along a chain of dependencies. We can therefore rigidly refer to fictional entities, as well as we refer to concrete entities. Actually, Kripke himself has held different positions regarding the use of fictional proper names. Whereas Kripke (1963) admits reference to merely possible entities, he retracts this thesis in *Naming and Necessity* (1972: 24, 158). On the one hand, Kripke (1972: 156) advocates a metaphysical impossibility. In a nutshell, any attempt to describe a possible world in which a fictional entity exists, say Holmes, would be partial. Beyond the fact that proper names are not synonymous with definite descriptions, and that they do not convey any essential property, various entities in different possible worlds might be referred to by the name "Holmes". As a consequence, nobody would be

¹ See e.g. Howell 2002) and Yagisawa 2001 on the problem of how to create fictional entities, Voltolini 2006 on the instant of creation, Everett 2005, 2013 and Friend 2007 on ontic vagueness and the difficulty of defining relevant identity criteria for created entities.

² See e.g. Fine 1995, Simons 1987, Correia 2005 for critical studies of ontological dependency. See also Fontaine and Rahman 2014 for solutions within a semantics for ontological dependencies in a modal framework.

suitably designated by “Holmes”. On the other hand, Kripke (1972: 157) holds an epistemological thesis. Even if someone corresponded to the description of Holmes in such a counterfactual situation, this would not suffice to decide whether it is or not the reference of “Holmes”. We would never be sure that Conan Doyle was actually referring to this man. According to Thomasson’s ATF, given that Conan Doyle was actually referring to an abstract artifact he created, we can use the name “Holmes” whose reference has been transmitted along a chain of dependencies, in particular a chain of copies. In the end, Thomasson meets the artifactual theory defended by Kripke (2013: 73), in which he finally considered fictional entities as creations we might refer to.

However, although Thomasson defines identity conditions for abstract artifacts,³ this only represents one side of the coin. Indeed, fictional characters are abstract artifacts outside the fiction, but they are (possibly) concrete entities inside the fiction. “[H]ow George Eliot, by imagining (or otherwise mentally engendering) a concrete and seemingly not-really-existent man, thereby creates the existent abstract non-man whom she baptizes ‘Silas Marner’”? can we ask with Howell (2002: 283). In the same vein, Yagisawa (2001: 158) finds “very difficult to fathom how Dickens could create an individual by writing a story which, when finished, would be a massively false story about that individual”. How, by describing a concrete human being, for example, with properties such as ‘being a weaver’, ‘being a man’, ‘being a Calvinist’, etc., an author would create an abstract artifact that has none of these properties? If an author creates a character by imaging or describing it, this character should have the properties that describe it in a way or another. Before addressing the question of the use of singular terms⁴ in fiction, we must therefore define fictional entities as double-aspect entities.⁵ We must explain

³ See Thomasson 1999: 63 for a sufficient criterion and Thomasson 1999: 67 for a necessary criterion of identity.

⁴ The notion of singular term covers the notions of proper name or individual constant, but also individual variables (free or bound).

⁵ The notion of *double-aspect* has its roots in Woods’s theory of fictionality, where it is introduced for slightly different reasons and without commitment to a fictional realism. See Woods and Isenberg 2010 and Woods 2018: 135 ff for more details.

how a fictional entity can take the appearance of an abstract artifact from an external viewpoint on fiction, but (possibly) the appearance of a concrete entity from an internal viewpoint. A fictional entity cannot be reduced to any of these aspects, which must be articulated in a unified theory.

In what follows, we first clarify the distinction and the articulation of the external and internal viewpoints by introducing a pair of fictionality operators at the propositional level, and we give their semantics in a modal framework (section 1). When passing to the first-order level, we need to define more precisely how we conceive fictional individuals in a modal framework. This we do by relying on Hintikka's world lines semantics, in which individuals are conceived as world lines which connect different world-bounded objects across the various possible worlds of a modal framework (section 2). How world lines are drawn is a complex matter we briefly discuss in relation to the question of identity of fictional characters (section 3). In this context, we introduce singular terms. We discuss their semantics with reference to problematic inference rules (i.e. the Substitution of Identicals and Existential Generalization) and the notion of creation (section 4). Our target is eventually to motivate a non-rigid interpretation of proper names in fiction compatible with the ATF, although cross-fictional reference (e.g. to real entities) is made possible by the interpretative efforts of the reader.

1 Fiction in a modal framework

Both viewpoints on fiction can be articulated in the language by means of a fictionality operator to be read as "according to the fiction". In its scope, we grasp the internal viewpoint. For example, "according to *Silas Marner: the Weaver of Raveloe*, Silas Marner is a weaver" is true, but not "according to *Silas Marner: the Weaver of Raveloe*, Silas Marner is a creation of George Eliott". By contrast, outside the scope of the fictionality operator, "Silas Marner is a creation of George Eliott" is true, but not "Silas Marner is a weaver". How to determine what is true in the fiction rapidly becomes a complex matter and drastic simplifications will be assumed here, even though sophistications of our basic approach might be considered later.

We introduce in the formal language a pair of fictionality

operators $[F]$ and $\langle F \rangle$. Their intended meaning is “according to the fiction...” and “it is compatible with the fiction...”, respectively.⁶ They will be given a modal interpretation, by considering that a fiction (partially) describes a possible world. A world *compatible* with the fiction is a world in which (at least) all that is said in the fiction is true. The *content* is the set of all logical consequences of a literary text. Sentences pertaining to the content are true in every world compatible with the fiction, and $[F]\varphi$ is true if φ forms part of the content. The literary text can be enriched by sentences that are not prescribed by the content, but are nevertheless compatible with the fiction, possibly on the basis of interpretive efforts of the readers. This yields various interpretations of the literary text, and $\langle F \rangle\varphi$ is true if φ is true in some but not all the worlds compatible with the fiction.⁷ In this respect, we are not ourselves committed to modal realism. A possible world is determined by a set of sentences. The machinery of possible worlds is a conceptual machinery useful for semantic purposes. When we talk in terms of ‘worlds’, ‘possible worlds’, or ‘worlds compatible with the fiction’, we do not mean something like ‘the world of the fiction’, as a reality distinct from ours. That is why the totality of interpretations considered may be narrower than what the linguistic meaning of the text alone would give rise to: this is because we allow the interpretative efforts of the reader to intervene, as it will be clear below. The impact of the reader can be modelled via a number of sentences added to the original text so as to determine the relevant interpretations as circumstances making true all the sentences of this larger text.

We now define a modal framework F like $F=(W, \mathcal{R})$ where W is a set of possible worlds w and \mathcal{R} is a set of accessibility relations between those worlds. For the moment, we only consider a relation R_F which relates to the worlds compatible with the relevant fiction \mathcal{F} .⁸

⁶ The fictionality operator has been introduced by Woods (1974) in order to distinguish explicitly the external and the internal viewpoints on fiction. A modal interpretation of the fictionality operator was proposed by Lewis (1978).

⁷ For more detailed definitions of the notions of content and interpretation in the context of the ATF, see Fontaine and Rahman 2010, the definitions of which are inspired by less formal definitions of Thomasson (1999, 64 ff).

⁸ Let $R_F(w)$ be the set of worlds R_F -accessible from w . A world w' is such that

At the propositional level, a model M consists of a frame F and a valuation v such that for every φ , $v(\varphi)$ is a subset of W in which φ is true (so that $M, w \models \varphi$ iff $w \in v(\varphi)$).⁹ The clauses for the introduction of the fictionality operators are defined as follows:

- (i) $M, w \models \langle \mathcal{F} \rangle \varphi$ iff for at least one $w' \in W$: $wR_{\mathcal{F}}w'$ and $M, w' \models \varphi$,
- (ii) $M, w \models [\mathcal{F}] \varphi$ iff for every $w' \in W$ such that $wR_{\mathcal{F}}w'$: $M, w' \models \varphi$,

An index can be added to the operator in order to explicit the relevant fiction (e.g. $[\mathcal{F}]_{\beta}$ if the relevant fiction is β). If w is the actual world and if $[\mathcal{F}]_{\text{Silas Marner}}$, then the situation discussed at the beginning of this section can be depicted as follows:

- (1) $M, w \models [\mathcal{F}]_{\text{Silas Marner}}$ Silas Marner is a weaver
- (2) $M, w \not\models [\mathcal{F}]_{\text{Silas Marner}}$ Silas Marner has been created by George Eliott
- (3) $M, w \models$ Silas Marner has been created by George Eliott
- (4) $M, w \not\models$ Silas Marner is a weaver

The interpretations are linked to the reader's perspective. The reader performs inferences from the content and extends the set of sentences true in the fiction. Thus, he reduces the set of worlds compatible with the fiction. For example, if it is not said in the fiction whether Silas Marner wears underpants, the readers can infer that he does, or not; so that:

- (5) $M, w \models \langle \mathcal{F} \rangle_{\text{Silas Marner}}$ Silas Marner wears underpants
- (6) $M, w \models \langle \mathcal{F} \rangle_{\text{Silas Marner}}$ Silas Marner does not wear underpants
- (7) $M, w \not\models [\mathcal{F}]_{\text{Silas Marner}}$ Silas Marner wears underpants

$w' \in R_{\mathcal{F}}(w)$ (in other words $wR_{\mathcal{F}}w'$) if w' is compatible with the fiction \mathcal{F} ; i.e. at least all the sentences of the content of \mathcal{F} are true in w' . In the context of the ATF, we might also add the restriction that a world w' can pertain to $R_{\mathcal{F}}(w)$ only if there exists a copy of \mathcal{F} in w .

⁹ Later, we will define models for first-order modal logic.

If we pay a peculiar attention to the real practices of the reader, and if the evaluation is made relative to a particular reader, it will not always be the case that either (5) or (6) be true. In the same vein, the reader cannot infer all the logical consequences of the content.¹⁰ This would demand unsustainable efforts. The point might be explained in relation to the distinction between *consequence-having* and *consequence-drawing* put forward by Woods in his logic of inference.¹¹ Consequence-having occurs in the logical space and is concerned with the entailment-relation, independently of the agents. Consequence-drawing occurs in the reasoner's mind, in the psychological sphere; that is, how the agent actually infers conclusions from a set of premises. Since an agent cannot draw conclusions that are not consequences of the set of premisses, consequence-drawing is not possible without consequence-having. But the agent cannot infer (and imagine) all the logical consequences of the content. Although our modal framework serves the purpose of consequence-having, it provides the conditions of possibility for consequence-drawing in terms of a plurality of interpretations.

That is why, in order to really grasp the reader's perspective, a logic of inference in fiction is required. In this paper, a very general and perhaps oversimplified understanding of the content and the definition of truth in fiction will be enough.¹² Of course, our semantics can be refined with rules of inferences applied in the construction of interpretations. For example, Friend (2017) argues in favour of the real foundations of the fictional worlds and the Reality Assumption, essential to our ability to understand fiction. Woods and Isenberg (2010) consider that real facts that are not untrue in the fiction can be imported as defeasibly true in the fiction, this is the Anti-Closed

¹⁰ This point can be understood in relation to the failure of closure under entailment and logical omniscience, but we cannot discuss in detail these points here.

¹¹ For the distinction between consequence-having and consequence-drawing, see Woods 2013: 24, 2018: 14.

¹² Even the content is difficult to define. Let us think about Conan Doyle's novels. If the narrator is not reliable, if Watson is sometimes lying (as it might be suggested by the text itself), the account of the content should be refined (this example was suggested by Stacie Friend, during the III Blasco Disputatio workshop held in Valencia (2017)).

World Assumption. Priest (2005: 89) too, allows an object to vary arbitrarily outside its determined properties in the limits of the constraints imposed by existing objects. As it stands, our semantics is compatible with weak versions of these principles. That is, they can be applied in the scope of the interpretation relative operator $\langle \mathcal{F} \rangle$, but not in the scope of the content relative operator $[\mathcal{F}]$. They can explain how interpretations are built, but they have no influence on the content. In some sense, we agree with Friend's and Woods's assumptions, but only at the level of interpretation.

2 Individuals in a modal framework

Before implementing quantification and singular terms in the language, we must introduce the notion of individual in the modal framework. When a language is interpreted with respect to a modal framework, which consists of various possible worlds, we have to think of individuals as entities that appear in different circumstances. Modal language, if it makes sense, thus presupposes a notion of individual which explains how it is possible to speak about the same entity in several possible worlds. What does it mean to say that an object X in a world w_1 is the same as an object Y in a world w_2 ? According to Tulenheimo (2017: 11–2), this question is meaningless if we understand identity in quantitative terms. Indeed, two objects in two different worlds are inevitably different. This is even more striking if we combine different viewpoints represented by different possible worlds in which X can be an abstract artifact and Y a concrete entity. Fictional characters, as double-aspect entities, presuppose a notion of individual that connects different objects across various possible worlds, but which is not reduced to any of their manifestations, or aspects.¹³

More precisely, given a modal framework, each possible world has its proper domain of (world-bounded) objects. In addition, a set of individuals, which manifest themselves in different possible worlds by taking the appearance of world-bounded objects, is presupposed. Individuals are not part of any world in particular. There is no mark

¹³ See Tulenheimo (2017: 44–51) for a comparison between this notion of modal individual with Kripke's and Lewis's approaches.

of them in the modal language either. They are nothing more than a precondition of first-order modal language.¹⁴ They are supposed to be these entities that have manifestations in various possible worlds. To put it in Hintikka's terms:

[I]n a context involving modal notions individuals have to be considered as members of several different possible worlds. An individual virtually becomes, for logical purposes, tantamount to the 'world line' [...] connecting its manifestations in these possible worlds. (Hintikka 1970a: 870)

When we say that an object X (e.g. an abstract artifact) in a world w_1 is the same as an object Y (e.g. a concrete human being) in a world w_2 , we mean that X and Y are linked by a world line. Stated otherwise, X and Y are the manifestations or the embodiments of the same individual in different possible worlds. As suggested by Hintikka (1970b: 412), world lines can be mathematically represented by individual functions, whose argument is a possible world w and whose value is an object of the domain of that world w .

Let F be a first-order modal framework such that $F=(W, \mathcal{R}, D, Q)$. W is a set of possible worlds w . \mathcal{R} is a set of accessibility relations between those worlds. We can distinguish between R_f that relates to the worlds compatible with a given fiction (internal viewpoint) and R_c that relates to the worlds compatible with creation (external viewpoint).¹⁵ We assume that each world $w \in W$ is associated with a set of (world-bounded) objects Q_w . D is a set of individual functions

¹⁴ See Hintikka and Sandu 1995, and Tulenheimo's (2017: 20) 'transcendental interpretation' of the world lines, on that point.

¹⁵ In the ATF, abstract artifacts are characterized by means of ontological dependence relations. These relations are intensional; i.e., they are defined over a plurality of worlds. For example, that Holmes historically rigidly depends on Conan Doyle means that in every world in which Holmes exists as an abstract artifact, Conan Doyle exists (or existed) as well, plus other constraints that have been defined in a modal-temporal framework by Fontaine and Rahman (2014). Worlds compatible with these conditions of existent for abstract artifacts are worlds compatible with creation. Let $R_c(w)$ be the set of worlds R_c -accessible from w . A world w' is such that $w' \in R_c(w)$ (in other words $w R_c w'$) if w' is compatible with the conditions of creation in w . Notice that $R_c(w)$ and $R_f(w)$ need not coincide, since in the worlds compatible with Conan Doyle's fiction Holmes does not appear as an abstract artifact.

(world lines), such that for every $d \in D$, $d(w) \in Q_w$, when they are defined (individual functions can be partial). The elements of a domain Q_w do not always coincide with the value of an individual function. We will call manifestation of an individual d in w an object q such that $q \in Q_w$ and $d(w) = q$.

The semantics is defined with respect to a model M obtained by adding to F an assignment g and an interpretation I .

[D1][ASSIGNMENT] An assignment M in is a function from the variables onto D . If g is an assignment defined on x , then $g(x)$ is an individual function. If this individual function is defined for w , the result $g(x)(w)$ of applying the function $g(x)$ to the world w is a local object belonging to Q_w . If g is an assignment and d is a world line, $g[x/d]$ stands for assignment that differs from g at most in that it assigns d to x .

The function I is a function from worlds and individual constants to $\bigcup Q_w$ for all $w \in W$, and from worlds and predicate symbols to $\bigcup Q_w^n$.¹⁶ For the moment, we define the interpretation of proper names non rigidly; that is, their reference may vary from one world to another.

[D2][VALUE OF A SINGULAR TERM] Value $\llbracket t \rrbracket_{M,w,g}$ of a term t in model M at world w under assignment g :

$$\llbracket t \rrbracket_{M,w,g} = I_w(t) \in Q_w, \text{ where } t \text{ is an individual constant,}$$

$$\llbracket t \rrbracket_{M,w,g} = g(t)(w) \in Q_w, \text{ where } t \text{ is a variable.}$$

[D3][INTERPRETATION OF A PREDICATE]

If P is an n -ary predicate, then $I(P,w) \subseteq Q_w^n$.

Let M be a model, the semantics is now defined as follows:

[D4][SEMANTICS]

$$(i) \quad M,w,g \models Pt_1, \dots, Pt_n \text{ iff } \langle \llbracket t_1 \rrbracket_{M,w,g}, \dots, \llbracket t_n \rrbracket_{M,w,g} \rangle \in I(P,w),$$

$$(ii) \quad M,w,g \models t_i = t_j \text{ iff } \llbracket t_i \rrbracket_{M,w,g}, \llbracket t_j \rrbracket_{M,w,g} \in Q_w \text{ and } \llbracket t_i \rrbracket_{M,w,g} = \llbracket t_j \rrbracket_{M,w,g},$$

¹⁶ The interpretation function may only be partial if we want to allow individual constants that have no referent in some worlds. Such details are not needed here.

- (iii) $M, w, g \models \neg\varphi$ iff $M, w, g \not\models \varphi$,
- (iv) $M, w, g \models \varphi \wedge \psi$ iff $M, w, g \models \varphi$ and $M, w, g \models \psi$,
- (v) $M, w, g \models \varphi \vee \psi$ iff $M, w, g \models \varphi$ or $M, w, g \models \psi$,
- (vi) $M, w, g \models \varphi \rightarrow \psi$ iff $M, w, g \models \neg\varphi$ or $M, w, g \models \psi$,
- (vii) $M, w, g \models \exists x\varphi$ iff for at least one $d \in D$: $d(w) \in Q_w$ and $M, w, g[x/d] \models \varphi$,
- (viii) $M, w, g \models \forall x\varphi$ iff for every $d \in D$ such that $d(w) \in Q_w$: $M, w, g[x/d] \models \varphi$,
- (ix) $M, w, g \models \diamond\varphi$ iff for at least one $w \in W$: wRw' and $M, w', g \models \varphi$,
- (x) $M, w, g \models \square\varphi$ iff for every $w \in W$ such that wRw' : $M, w', g \models \varphi$.

The semantics is completed by adding the clauses for the fictionality operators (see section 1). In the clauses (ix) and (x), the modal operators can be replaced by other operators, possibly together with other kinds of relevant accessibility relations.¹⁷

In world lines semantics, individuals (i.e. those entities that can be the value of a bound variable) are individual functions (or world lines) and not world-bound objects. That is why, in the clauses (vii) and (viii), the quantifiers range over D and not Q_w . One restriction is that they range over locally manifested individuals; that is, in a world w , we quantify over the individuals d such that $d(w) \in Q_w$. For example, $M, w, g \models [F]\exists xPx$ iff for every w' such that $wR_{\mathcal{F}}w'$, there is a $d \in D$ such that $d(w') \in Q_{w'}$ and $M, w', g[x/d] \models Px$. Then, in accordance with (i), $M, w', g[x/d] \models Px$ iff the value of the singular term x in w' under g satisfies Px in w' . In other words, $M, w', g[x/d] \models Px$ iff the (world-bound) object $d(w') \in Q_{w'}$ is such that $d(w) \in I(P, w')$. By contrast, $M, w, g \models \exists x[F]Px$ iff there is a $d \in D$ such that $d(w) \in Q_w$, and $M, w', g[x/d] \models Px$ for every w' such that $wR_{\mathcal{F}}w'$. Then, $M, w', g[x/d] \models Px$ iff that $d(w) \in I(P, w')$. Given that $I(P, w') \subseteq Q_w$, $g(d)(w') \in Q_{w'}$. As a consequence, whereas $\exists x[F]Px$ assumes the apparition of d in w and w' as well, $[F]\exists xPx$ does not.

¹⁷ For example, the operator of necessity can be replaced by a belief operator \mathcal{B} and the accessibility relation replaced by $R_{\mathcal{B}}$ that relates to the worlds compatible with the belief state of an agent.

3 Drawing world lines

How the world lines are drawn is a very complex matter. In general, they are drawn by an agent, depending on the context, a modality, and sometimes by assuming criteria of cross-identification; that is, criteria that contextually allow the agent to re-identify an individual on various occasions. As stressed by Hintikka (1969: 109–10) in a very Kantian way, they are themselves human artifacts, which depend on modes of identification, have an objective reality, and make possible our transaction with reality. Such a Kantian conception of individuals is also a reason why world lines semantics is well suited for the ATF: fictional individuals are human mind constructions. The fact that modes and criteria of identification are not infallible, and that the re-identification processes might fail on some occasions, explains why the world lines might split or merge, and why identity is contingent. Whereas Hintikka has distinguished between two modes of identification,¹⁸ the following definition provides a criterion of individuation for literary fictional characters in the context of their creation:¹⁹

[D5][LITERARY FICTIONAL INDIVIDUAL] An individual *d* is a *literary fictional individual* iff it satisfies the two following conditions:

- (i) it is an abstract artifact satisfying the relevant ontological dependencies in the worlds compatible with creation (external aspect),
- (ii) it has the characterizing properties in the worlds (partially) described by the relevant story (internal aspect).

According to this definition, a literary fictional individual is defined according to its two aspects. On the one hand, it must satisfy the relevant ontological dependencies in the worlds compatible with creation; that is, worlds in which its manifestations are tight to the objects on which it depends. In the worlds compatible with creation,

¹⁸ See Hintikka 1970a: 873 and Hintikka and Sandu 1995: 274 ff on the perspectival and public modes of identification.

¹⁹ This definition is an adaptation of the definition given by Fontaine and Rahman (2014: 513), who also define the relevant ontological dependencies mentioned in [D5] (i).

Holmes can have a manifestation only if Conan Doyle has created it (rigid historical dependence) and only if there are copies of the original work and/or a competent readership (generic constant dependence). In these worlds, Holmes manifests itself as an abstract artifact. On the other hand, it must satisfy the properties he is characterized as having in the worlds compatible with the fiction (in which it arises for the first time); i.e. all the properties that logically follow from the meaning of the text. In those worlds, Holmes is a human being, a detective, etc., but he is not an abstract artifact (and he has not been created by Conan Doyle). We can distinguish the worlds compatible with creation and those compatible with fiction by means of different accessibility relations (as previously suggested).

The definition above yields a syncretistic account of fictional individuals, by considering the context of creation and the properties of the fictional character. According to Fontaine and Rahman (2014: 509), creation is linked to an achieved codification act. Codification is a “completed linguistic act (or process) that has been made public”. The author fixes a codex or a canon, by means of which some constructive instructions are given for the (re-)construction of the fictional characters and their properties, in a way similar to instructions given to construct a proof or a canonical object in Constructive Type Theory.²⁰ The codification act yields a constructive process type that can be repeated, whenever an existing copy or a reader’s memory sustains the existence of such a set of instructions. This proposal is compatible with Voltolini (2012: 563) who combines a set-theoretical element (set of properties) with a game-theoretical one (a make-believe element).²¹ In both cases, a syncretistic definition assumes a stronger link between the fictional character and its internal properties. As a consequence, different creative acts will produce different fictional individuals. This is consistent with the ATF, but how is it possible to account for cross-fictional identity?

Notice that the criteria defined in [D5] explains how the world

²⁰ More details on Constructive Type Theory can be found in Martin-Löf 1984. A more recent introduction to Constructive Type Theory, including the notions of e.g. proof-objects and canonical objects can be found in Rahman et al. 2018.

²¹ In that case, the set of instructions is not relative to a codification, but rather to the rules of a game.

line is drawn, but only in the context of creation; i.e. in relation to the original work in which the literary fictional character is created. It does not prevent us from extending its corresponding world line across the worlds compatible with other fictions, or even other kinds of worlds (compatible with the representations, beliefs, etc., of intentional agents), on the basis of other criteria. However, this is always a matter of interpretation. For example, Lewis Carroll seems to have created Alice across various works; namely *Alice's Adventures Underground* and *Alice in wonderland*. But each work is the result of different codification acts, and a new fictional individual with its own properties is created by each of them. Then, how the different stories could be about the same character (Alice)? That question is answered at different levels. Strictly speaking, the two works yield two different fictional individuals—say d_1 and d_2 respectively. However, when reading *Alice in wonderland*, nothing prevents the reader from interpreting that work by considering that both individuals are identical. That is, even if the content does not prescribe the manifestation of d_1 in the worlds compatible with *Alice in wonderland*, it does not prevent its apparition in some (but not all) those worlds. In other words, let the identity between d_1 and d_2 be a relation that is not determined by the fiction. By his interpretative efforts, the readers can infer their identity by extending the world line of d_1 across some (but not all) worlds compatible with the story. This can be done on the basis of contextual, descriptive or other kinds of identification criteria.

More formally, by taking into account the interpretative efforts of the readers, it can be the case that $d_1(w)=d_2(w)$ for some (but not all) worlds w compatible with the story. Given a relevant model M , the situation might be described as follows:

$$(8) \quad M, w, g [x/d_1, y/d_2] \not\models [F]_{Alice\ in\ Wonderland} (x=y \wedge x=Alice)$$

$$(9) \quad M, w, g [x/d_1, y/d_2] \models \langle F \rangle_{Alice\ in\ Wonderland} (x=y \wedge x=Alice)$$

Let us recall that, according to [D4], clause (ii), identity is defined over world-bounded objects. Moreover, since the criteria of re-identification are fallible, world lines can merge and split. According to (9), there are some worlds compatible with *Alice in Wonderland*, given determinate interpretations, in which the two world lines merge and share the same manifestation, a world-bounded object called “Alice”

in those worlds. Although we do not advocate for a descriptivism (i.e. that proper names are synonymous with a definite description), such identity can be inferred on the basis of descriptive similarities or even some kind of descriptive content conveyed by the proper name “Alice”. In general, identity between the manifestations of different fictional individuals is contingent; that is, even if $d_1(w_i)=d_2(w_i)$ in a w_i such that $wR_{\neq}w_i$ for two distinct individuals d_1 and d_2 , there can be a w_j such that $wR_{\neq}w_j$ in which $d_1(w_j)\neq d_2(w_j)$.²²

To conclude, we might expect a weaker link to the codification act, and world lines drawn on the basis of pragmatic criteria. From the perspective of a cultural community, it might be relevant to consider, not the contexts of creation of each fictional character separately, but a wider context in which a larger world line would connect all the manifestations of both individuals. This often happens within literary circles, in which different characters are gathered together in a “bigger” fictional character. Nonetheless, as stressed by Voltolini (2006) what is needed here is a relation weaker than identity. Both characters are not identical; they form part of a bigger entity. Our semantics accounts for this phenomenon as follows. Let d_1 and d_2 be the two respective Alices of *Alice’s Adventures Underground* and *Alice in wonderland*. Then, let d_3 be a wider world line, drawn for pragmatic or interpretive reasons. This might be an individual we actually talk about when we talk about Alice without making reference to any work in particular. In every world w compatible with *Alice’s Adventures Underground*, $d_1(w)=d_3(w)$. And in every world w' compatible with *Alice in wonderland*, $d_2(w')=d_3(w')$. It can even be the case that both coincide in the actual world. But, in the worlds compatible with *Alice’s Adventures Underground* and *Alice in wonderland*, the world lines d_1 and d_2 behave independently one of another. Strictly speaking, they are not identical world lines, since they could split for other reasons beyond the worlds in which their manifestations coincide.²³

²² By contrast, in a standard framework (without world lines), we quantify over ‘cross-world’ objects, which form part of the domain of different worlds. Identity is a relation between such objects and is thus necessary. According to Kripke (1980), since proper names are interpreted rigidly, $k_1=k_2\rightarrow\Box k_1=k_2$ is valid, as well as $\forall x\forall y(x=y\rightarrow\Box x=y)$ according to Fitting and Mendelsohn (1998: 146 ff).

²³ I am thankful to Manuel García-Carpintero for having suggested the

4 Proper names

Up to now, we have interpreted proper names non rigidly. As in the previous example, “Alice” can refer to the manifestations of different individuals depending on the world in which it is interpreted. It can also be used to refer to world-bounded objects that are the manifestations of different fictional individuals. Things are more striking when we think of the use of “Napoleon” in Orwell’s *Animal Farm*, for example. In the worlds compatible with *Animal Farm*, “Napoleon” refers to the manifestations of a fictional individual created by Orwell. Literary critics have emitted divergent interpretations of that work. Some say that “Napoleon” might refer to the manifestations of the historical character Stalin, or Napoleon, or both, depending on how we read the story.

In our view, the same holds in historical novels like Tolstoy’s *War and Peace*, in which the identity between the historical and the fictional Napoleon is a matter of interpretation. It is true that the use of the name “Napoleon”, either in *Animal Farm* or in *War and Peace*, might be used by the authors as an invitation to imagine particular facts about the historical Napoleon. But, this forms part of the interpretation, not the content. Although the use of non-rigid proper names can be associated with a descriptive content, we need not consider that they are synonymous with a definite description either, and such a descriptive content would also be relative to a context, an agent, and a lot of other criteria.²⁴ Moreover, criteria of identification cannot be expressed in the modal language. Indeed, in order to express these criteria, we should quantify in intensional contexts, across possible worlds. But quantifying in intensional context presupposes that world lines have already been drawn, and that criteria of identification have already been given. Otherwise, what would we quantify on? This would be doomed to circularity. According to Hintikka and Sandu (1995: 249), this is due to the transcendental nature of the question. We

necessity of taking into account pragmatic considerations during the III Blasco Disputatio workshop held in Valencia (2017).

²⁴ See García-Carpintero 2015 for a descriptivist account of the invitation to imagine facts concerning a real entity.

can only presuppose that criteria of identification have been given, we cannot express them. In fiction, as well as in other intensional contexts, the system of references and the system of individuals are mutually independent.²⁵

The interpretative-relative use of proper names in fiction we defend here is a corollary of the weakened application of Friend's Reality Assumption and Woods's Anti-Closed Worlds Assumption we have previously discussed (see section 1). When reading a fiction, readers will generally understand real proper names as referring to historical or real entities unless there are contrary indications in the text. From a pragmatic perspective, we also recognize that a reader needs this kind of assumption to actually understand the fiction, otherwise the interpretative efforts would become unsustainable. However, we need not consider that a name like "Napoleon" rigidly refers to the historical Napoleon. We only need to assume that the world lines of different individuals can be extended and merge in some (but not all) worlds compatible with the fiction, depending on the interpretation. And this is precisely what our modal approach makes possible. Literary critics might object that a *correct* interpretation of *War and Peace* presupposes that "Napoleon" refers to the historical Napoleon. We answer that we should not confuse the set of *possible interpretations* with the set of *correct interpretations*, the latter assuming certain rules of inferences based upon a certain cultural and literary background.

What about rigidity in the world lines semantics? Technically, it is possible to interpret proper names rigidly. If we do so, the interpretation must be defined over the set of individuals (the world lines). It might seem unnatural to think of the interpretation of proper names in terms of entities that lie outside the possible worlds. However, since the objects are world-bounded entities, referring rigidly to them would not make sense. In order to interpret rigidly proper names, adjustments are needed in the general semantics (see section 2). First, the function I is now a function from individual constants to D ; i.e. $I(k) \in D$, the same individual no matter the world relatively to which k is interpreted. Then, the value of a singular term must now be defined as follows:

²⁵ See Hintikka and Hintikka 1989: 159–60.

[D2][VALUE OF A SINGULAR TERM] Value $\llbracket t \rrbracket_{M,w,g}$ of term t in model M at world w under assignment g :

$\llbracket t \rrbracket_{M,w,g} = g(I(t),w) \in Q_w$, where t is an individual constant,

$\llbracket t \rrbracket_{M,w,g} = g(t)(w) \in Q_w$, where t is a variable.

In the first clause, $I(t)$ is an individual $d \in D$, and $g(I(t),w)$ is the value of d in w .

For example, if “Napoleon” is a rigid designator, it refers to an individual, say d_1 , the same in every possible world, whether it be the actual world, a world compatible with *Animal Farm* or a world compatible with *War and Peace*. Thus, in a given world, the name “Napoleon” could not refer to the manifestations of different individuals, for example both the manifestation of the historical Stalin and the manifestation of Napoleon. In order to handle rigidity across the different possible interpretations, we should consider that there are two homonyms “Napoleon”, one that refers to the character created by Orwell and the other to the historical one. Let us introduce a name “Napoleon*” to disambiguate. We define a model for *Animal Farm*, with the following interpretation of proper names:

$$I(\text{Napoleon}) = d_1$$

$$I(\text{Napoleon}^*) = d_2$$

$$I(\text{Stalin}) = d_3$$

Notice that the interpretation is not world-relative anymore: it is defined over the set D of individuals. What about the following set of sentences? Is it consistent?

$$(10) \langle \mathcal{F} \rangle_{\text{Animal Farm}} (\text{Stalin} = \text{Napoleon})$$

$$(11) \langle \mathcal{F} \rangle_{\text{Animal Farm}} (\text{Stalin} = \text{Napoleon}^*)$$

$$(12) \langle \mathcal{F} \rangle_{\text{Animal Farm}} (\text{Napoleon} = \text{Napoleon}^*)$$

Unlike Kripkean semantics, necessary identity is not a consequence of rigidity in world lines semantics. It requires the extra-assumption

that world lines do not split and merge.²⁶ Since we do not make this assumption (see section 3), nothing precludes the possibility of having $d_1(w)=d_2(w)$ but $d_1(w')\neq d_2(w')$ where $d_1, d_2 \in D$ and $w, w' \in W$. More generally, no matter how proper names are interpreted, the well-known necessary identities and the Substitution of Identicals do not hold:

$$(13) \neq \forall x \forall y (x=y \rightarrow \Box x=y)$$

$$(14) \neq k_1=k_2 \rightarrow \Box k_1=k_2$$

$$(15) Pk_1, k_1=k_2 \neq Pk_2$$

Therefore, whether proper names should be interpreted rigidly or not cannot be decided on the basis of the behaviour of identity and the failure of the Substitution of Identicals in fictional or any other kind of intensional context.²⁷ Things go otherwise with Existential Generalization, whose validity assumes a uniqueness presupposition.²⁸ Indeed, if we admit the validity of the following inference rule:

$$(16) Pk_1 \models \exists x Px$$

then, we must assume that every proper name is associated with a unique reference (an individual) or the manifestations of a unique individual. In that case, the mutual independence of the system of individuals and the system of references would collapse. Indeed, each proper name would be associated with a well-identified individual. Therefore, the fundamental difference is the following: If proper names are rigidly interpreted, we are committed to the validity of Existential Generalization. If we reject its validity, we must reject rigidity.

²⁶ See Hintikka and Sandu 1995: 270.

²⁷ For the same reasons, Everett's (2005) paradox of *Frackworld* is easily solved, no matter the way proper names are interpreted. The different interpretations (with Frick = Frack and Frick \neq Frack) are possible only if we recognize the existence of two fictional individuals (Frick and Frack) whose respective world lines merge and split. And this does not involve any kind of ontic indeterminacy.

²⁸ See Fontaine 2019 for a more detailed study of that point.

First, the application of the principle of individuation [D5] might suggest that we interpret fictional proper names rigidly. Indeed, regarding *A Study in Scarlet*, the name “Holmes” refers to the manifestations of that individual created by Conan Doyle in the actual world, and in every world compatible with the story as well. We are thus committed with a kind of presupposition of uniqueness of reference, which at first sight commits to the validity of Existential Generalization. However, this is only a local presupposition that should not be confused with rigidity. Therefore, this case does not commit us to the validity of Existential Generalization.

Second, should (and can) we reject Existential Generalization (and consequently rigidity)? If Existential Generalization is valid, everything that can be named in fiction must be associated with a well-identified individual, created by the author. This might be in accordance with Thomasson’s ATF.²⁹ But this is a too strong version of creationism. Indeed, how an author could be able to create everything that appears in the fiction?³⁰ It seems that the position holding that everything in the fiction is created by the author is not tenable. And rejecting this position commits to the rejection of Existential Generalization; i.e. $[F]Pk$ can be true without having $\exists x[F]Px$. So, rigidity must also be rejected. In that sense, we argue for a weaker version of ATF according to which there are (world-bounded) fictional objects that are not well-identified literary fictional characters (individuals). We can refer to them, non rigidly, in a more or less vague manner. They cannot be properly re-identified in the different worlds compatible with the fiction or across the interpretations of different readers. In some sense, they are purely intentional objects, depending on the intentional states of a reader, or the author, and determinate interpretations of the work. For example, we could name several Uruk-hai in of *The Lords of the Rings*, without being able to identify them precisely and even if they do not correspond to abstract artifacts in the actual worlds.³¹

²⁹ See Thomasson 1999: 13, 88.

³⁰ It is not even sure that their identity could be defined (see e.g. Friend 2007 and Everett 2013: 191).

³¹ See Voltolini 2006: 209 on this example.

For the sake of comparison, it is interesting to highlight that Meinongian approaches to fiction are not affected by these issues in the same way. They are not pressed by the difficulty of creation: fictional entities are only characterized, and an object corresponds to every characterization. Therefore, intentionality is always directed to a well-defined entity, whether it exists or not. That is why the Meinongian logic of intentionality of Priest (2005) accommodates perfectly well the validity of Existential Generalization. It is even a requisite of his logic. In this context, a rigid interpretation of proper names is probably more relevant.

5 Conclusion

Several criticisms raised against Thomasson's ATF can be overcome by articulating both viewpoints on fiction. This we do by introducing fictional operators in the formal language. By giving them a modal interpretation, we define fictional characters as double-aspect entities. Therefore, when we quantify in fictional contexts, we quantify over individuals which manifest themselves as abstract artifacts in the worlds compatible with creation and (possibly) as concrete entities in the worlds compatible with the fiction. The world line of a literary fictional character connects abstract artifacts and concrete entities. Given the criteria of individuation [D5], we agree with a local presupposition of uniqueness of reference for fictional proper names. However, by insisting on the interpretative role of the reader, we have defended a non-rigid interpretation of proper names that allows a weakened version of the ATF, more viable with respect to the possibility of creation itself. Indeed, there are cases in which the codification does not provide sufficient instructions to reconstruct the apparition of an individual, in the same way that sometimes we do not have any means to re-identify the apparitions of the same individual under different circumstances. Of course, the definition of a well-identified individual can remain vague in some cases, but it seems that such a limitation is necessary to the viability of the ATF.³²

³² I acknowledge financial support of the postdoctoral program of the FCT-Portugal (Grant Number SFRH/BPD/116494/2016) and the FCT CFCUL UID/FIL/00678/2019. I warmly thank the organizers, Josep Corbí and Jordi Valor, and

Matthieu Fontaine
 Center for Philosophy of Sciences of the University of Lisbon
 Campo Grande, Edifício C4, Piso 3
 1749-016 Lisboa, Portugal
 fontaine.matthieu@gmail.com

References

- Correia, Fabrice. 2005. *Existential Dependence and Cognate Notions*. Munich: Philosophia Verlag.
- Everett, Anthony. 2005. Against fictional realism. *Journal of Philosophy* 102(12): 624–49.
- Everett, Anthony. 2013 *The Nonexistent*. Oxford: Oxford University Press.
- Fine, Kit. 1995. Ontological dependence. *Proceedings of the Aristotelian Society* 95: 269–90.
- Fitting, Melvin; and Mendelsohn, Richard. 1998. *First-Order Modal Logic*. Dordrecht: Kluwer.
- Fontaine, Matthieu. 2019. Hintikka, free logician—singular terms in world lines semantic. *Logica Universalis* 13: 179–201.
- Fontaine, Matthieu; and Rahman, Shahid. 2010. Fiction, creation and fictionality. *Methodos* 10: <http://journals.openedition.org/methodos/2343>.
- Fontaine, Matthieu; and Rahman, Shahid. 2014. Towards a semantics for the artifactual theory of fiction and beyond. *Synthese* 191(3): 499–516.
- Friend, Stacie. 2007. Fictional characters. *The Philosophy Compass* 2(2): 141–56.
- Friend, Stacie. 2017. The real foundation of fictional worlds. *Australasian Journal of Philosophy* 95(1): 29–42.
- García-Carpintero, Manuel. 2015. Is fictional reference rigid? *Organon F* 22 (Supplementary Issue 1): 145–68.
- Hintikka, Jaakko. 1969. *Models for Modalities*. Dordrecht: Reidel.
- Hintikka, Jaakko. 1970a. Objects of knowledge and belief: acquaintances and public figures. *Journal of Philosophy* 67(21): 869–83.
- Hintikka, Jaakko. 1970b. The semantics of modal notions and the indeterminacy of ontology. *Synthese* 21(3): 408–24.
- Hintikka, Jaakko; and Hintikka, Merrill. 1989. *The Logic of Epistemology and the Epistemology of Logic*. Dordrecht: Kluwer.
- Hintikka, Jaakko; and Sandu, Gabriel. 1995. The fallacies of the new theory of reference. *Synthese* 104(2): 245–83.
- Howell, Robert. 2002. Review of *Fiction and Metaphysics* by Amie L. Thomasson.

the participants of the III Blasco Disputatio Workshop (Valencia, Spain, 2017). Particular thanks to an anonymous reviewer who considerably helped me to improve my paper. I am also particularly thankful to Olga Pombo, Shahid Rahman and John Woods, among many others, for their support, comments, and fruitful advices.

- The Philosophical Quarterly* 52(207): 282–4.
- Priest, Graham. 2005. *Towards Non-Being. The Logic and Metaphysics of Intentionality*. Oxford: Clarendon Press.
- Kripke, Saul. 1963. Semantical considerations of modal logic. *Acta Philosophica Fennica* 16: 83–94.
- Kripke, Saul. 1972. *Naming and Necessity*. Cambridge: Harvard University Press.
- Kripke, Saul. 2013. *Reference and Existence: The John Locke Lectures*. Reprint with emendations of six lectures given at Oxford University between October 30th and December 4th, 1973. New York: Oxford University Press.
- Lewis, David. 1978. Truth in fiction. *American Philosophical Quarterly* 15: 37–46.
- Martin-Löf, Per. 1984. *Intuitionistic Type Theory: Notes by Giovanni Sambin of a series of lectures given in Padua, June 1980*. Naples: Bibliopolis.
- Rahman, Shahid; McConaughey, Zoe; Klev, Ansten; and Clerbout, Nicolas. 2018. *Immanent Reasoning of Equality in Action—A Plaidoyer for the Play Level*. Cham: Springer.
- Simons, Peter. 1987. *Parts: A Study in Ontology*. Oxford: Clarendon Press.
- Thomasson, Amie. 1999. *Fiction and Metaphysics*. Cambridge: Cambridge University Press.
- Tulenheimo, Tero. 2017. *Objects and Modalities—A Study in the Semantics of Modal Logic*. Dordrecht: Springer.
- Voltolini, Alberto. 2006. *How Ficta Follow Fiction: A Syncretistic Account of Fictional Entities*. Dordrecht: Springer.
- Voltolini, Alberto. 2012. Crossworks ‘identity’ and intrawork* identity of a fictional character. *Revue Internationale de Philosophie* 4(262): 561–76.
- Yagisawa, Takashi. 2001. Against creationism in fiction. *Noûs* 35(15): 153–72.
- Woods, John. 1974. *The Logic of Fiction: A Philosophical Sounding of Deviant Logic*. The Hague: Mouton.
- Woods, John. 2013. *Errors of Reasoning: Naturalizing the Logic of Inference*. London: College Publications.
- Woods, John. 2018. *Truth in Fiction—Rethinking its Logic*. Cham: Springer.
- Woods, John; and Isenberg, Jillian. 2010. Psychologizing the semantics of fiction. *Methodos* 10: <http://methodos.revues.org/2387>.