A **Meaningful** Road to **Explanation**

BIR 2024
Prague, Czech Republic

Giancarlo Guizzardi

UNIVERSITEIT TWENTE.

Semantics
Cybersecurity
Services

Explanation, **Semantics**, and $\binom{o}{o}$**ntology**

Contents lists available at ScienceDirect

Data & Knowledge Engineering

journal homepage: www.elsevier.com/locate/datak

Explanation, semantics, and ontology

Giancarlo Guizzardi [a,*], Nicola Guarino [b]

[a] Semantics, Cybersecurity & Services, University of Twente, Enschede, The Netherlands
[b] ISTC-CNR Laboratory of Applied Ontology (LOA), Trento, Italy

ARTICLE INFO

Keywords:
Real-world semantics
Ontology
Explanation
Ontological unpacking
Semantic interoperability

ABSTRACT

The terms 'semantics' and 'ontology' are increasingly appearing together with 'explanation', not only in the scientific literature, but also in everyday social interactions, in particular, within organizations. Ontologies have been shown to play a key role in supporting the semantic interoperability of data and knowledge representation structures used by information systems. With the proliferation of applications of Artificial Intelligence (AI) in different settings and the increasing need to guarantee their explainability (but also their interoperability) in critical contexts, the term 'explanation' has also become part of the scientific and technical jargon of modern information systems engineering. However, all of these terms are also significantly overloaded. In this paper, we address several interpretations of these notions, with an emphasis on their strong connection. Specifically, we discuss a notion of explanation termed *ontological unpacking*, which aims at explaining symbolic domain descriptions (e.g., conceptual models, knowledge graphs, logical specifications) by revealing their *ontological commitment* in terms of their so-called *truthmakers*, i.e., the entities in one's ontology that are responsible for the truth of a description. To illustrate this methodology, we employ an ontological theory of relations to explain a symbolic model encoded in the *de facto* standard modeling language UML. We also discuss the essential role played by ontology-driven conceptual models (resulting from this form of explanation processes) in supporting semantic interoperability tasks. Furthermore, we revisit a proposal for quality criteria for explanations from philosophy of science to assess our approach. Finally, we discuss the relation between ontological unpacking and other forms of explanation in philosophy and science, as well as in the subarea of Artificial Intelligence known as Explainable AI (XAI).

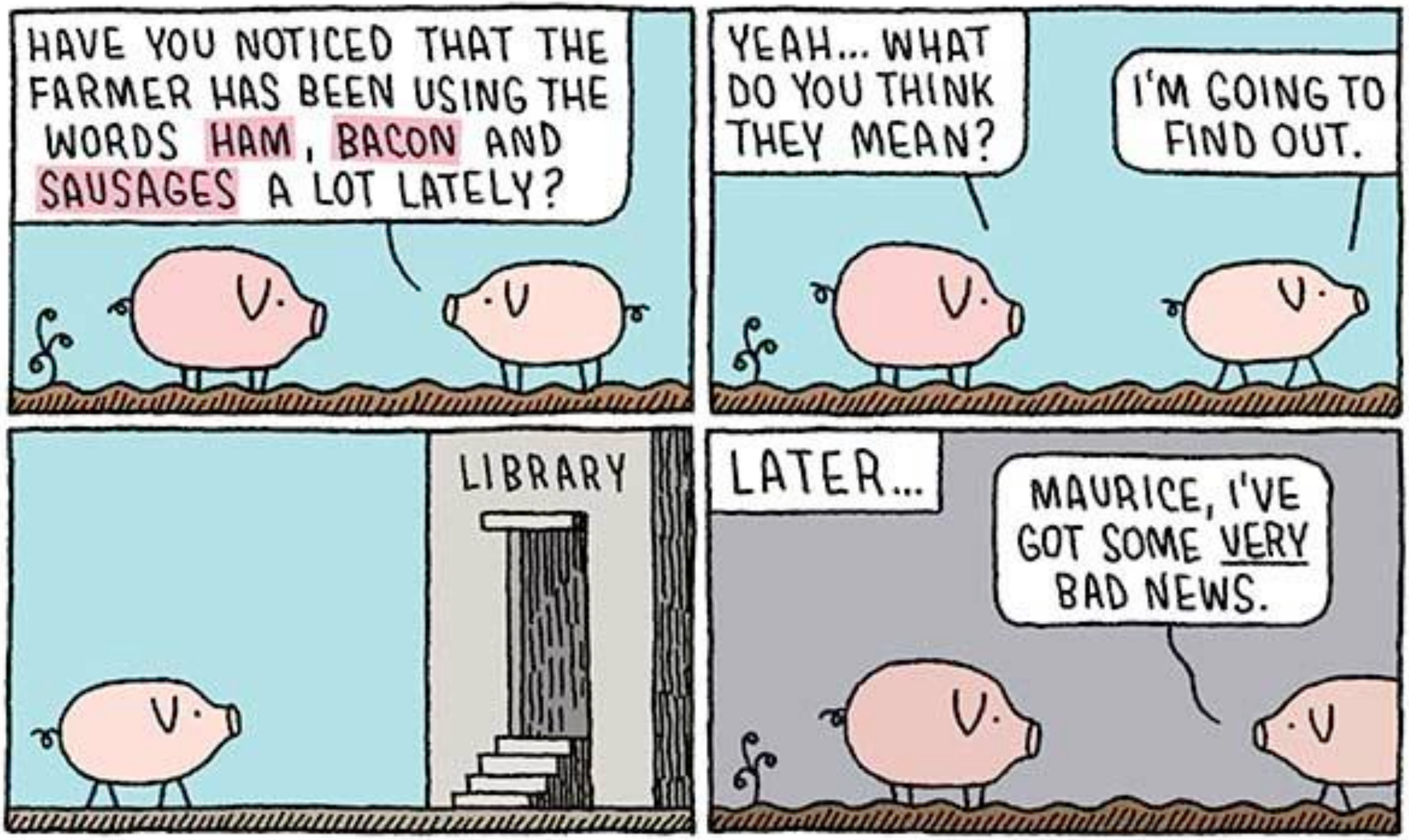
On the Multiple Roles of Ontologies in Explanations for Neuro-symbolic AI

Roberto Confalonieri [a,*] and Giancarlo Guizzardi [b]

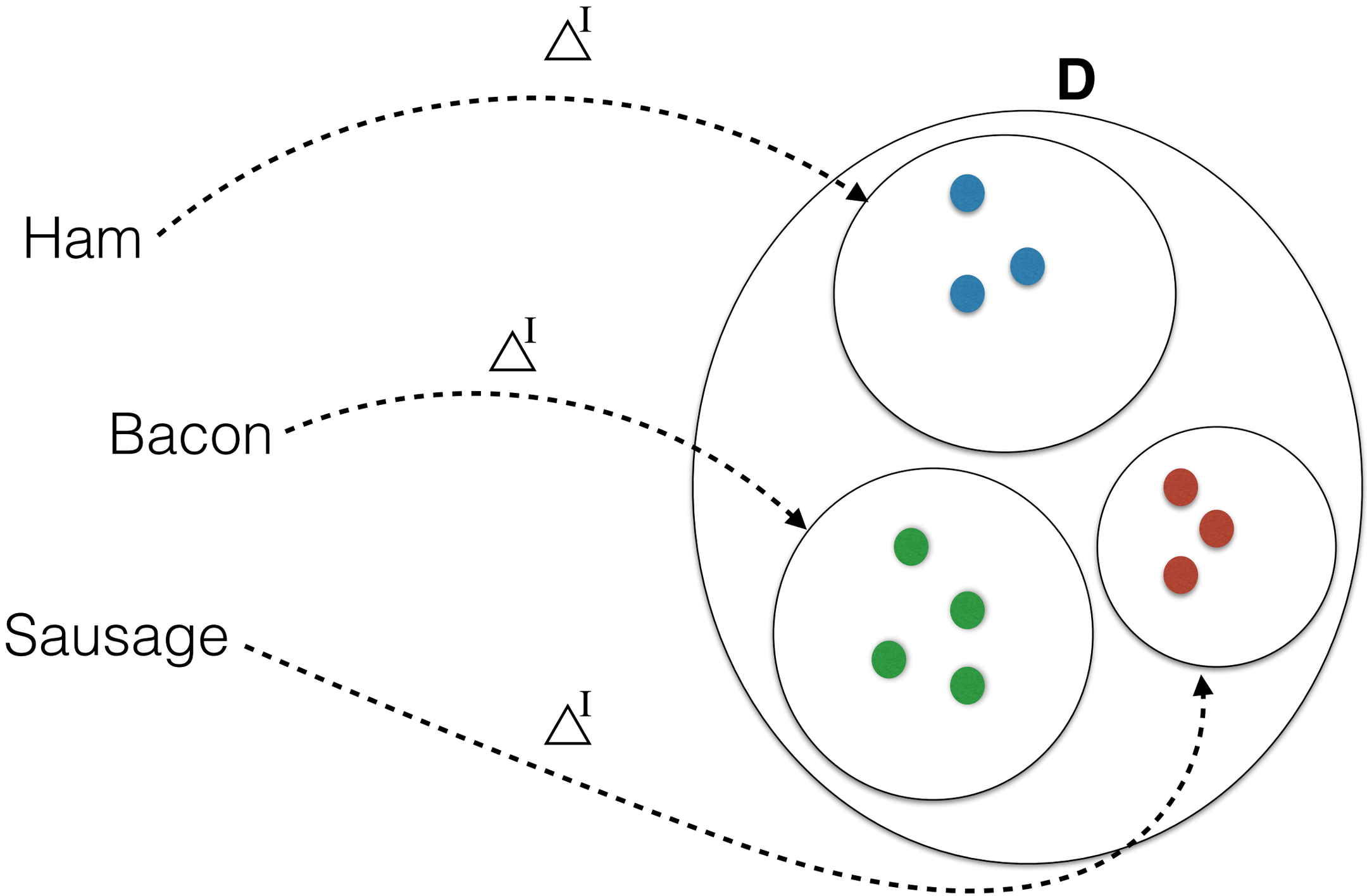
[a] *Department of Mathematics 'Tullio Levi-Civita', University of Padua, Padova, Italy*
E-mail: roberto.confalonieri@unipd.it
[b] *Department of Semantics, Cybersecurity & Services (SCS), University of Twente, Enschede, The Netherlands*
E-mail: g.guizzardi@utwente.nl

Abstract. There has been a renewed interest in symbolic AI in recent years. Symbolic AI is indeed one of the key enabling technologies for the development of neuro-symbolic AI systems, as it can mitigate the limited capabilities of black box deep learning models to perform reasoning and provide support for explanations. This paper discusses the different roles that explicit knowledge, in particular ontologies, can play in drawing intelligible explanations in neuro-symbolic AI. We consider three main perspectives in which ontologies can contribute significantly, namely reference modelling, common-sense reasoning, and knowledge refinement and complexity management. We overview some of the existing approaches in the literature, and we position them according to these three proposed perspectives. The paper concludes by discussing some open challenges related to the adoption of ontologies in explanations.

Keywords: Neuro-symbolic AI, Explanations, Applied Ontologies


HAVE YOU NOTICED THAT THE FARMER HAS BEEN USING THE WORDS HAM, BACON AND SAUSAGES A LOT LATELY?
YEAH... WHAT DO YOU THINK THEY MEAN?
I'M GOING TO FIND OUT.
LIBRARY
LATER...
MAURICE, I'VE GOT SOME VERY BAD NEWS.

by Tom Gauld


Ham
Bacon
Sausage
I
I
I
D

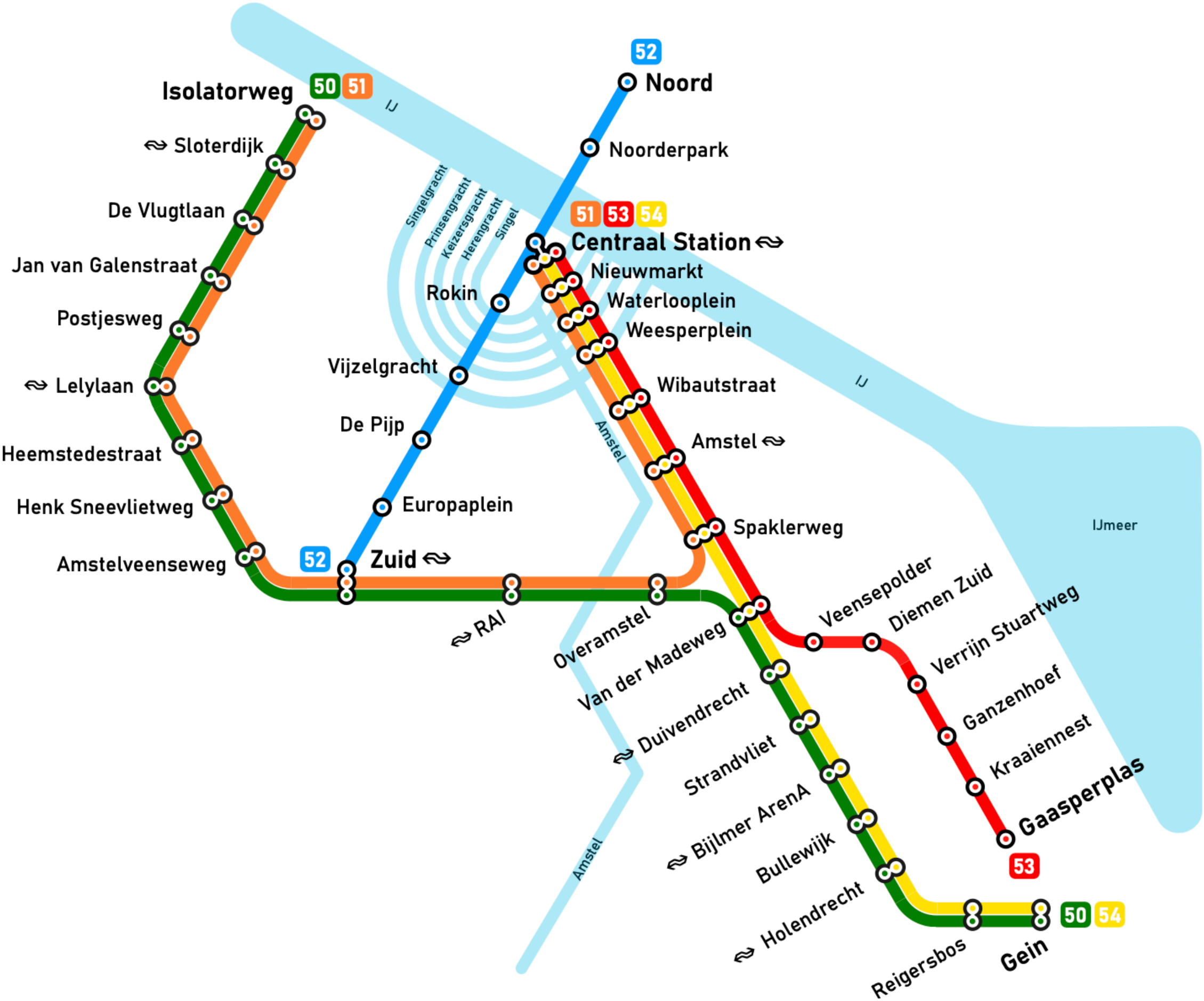
Formal Semantics

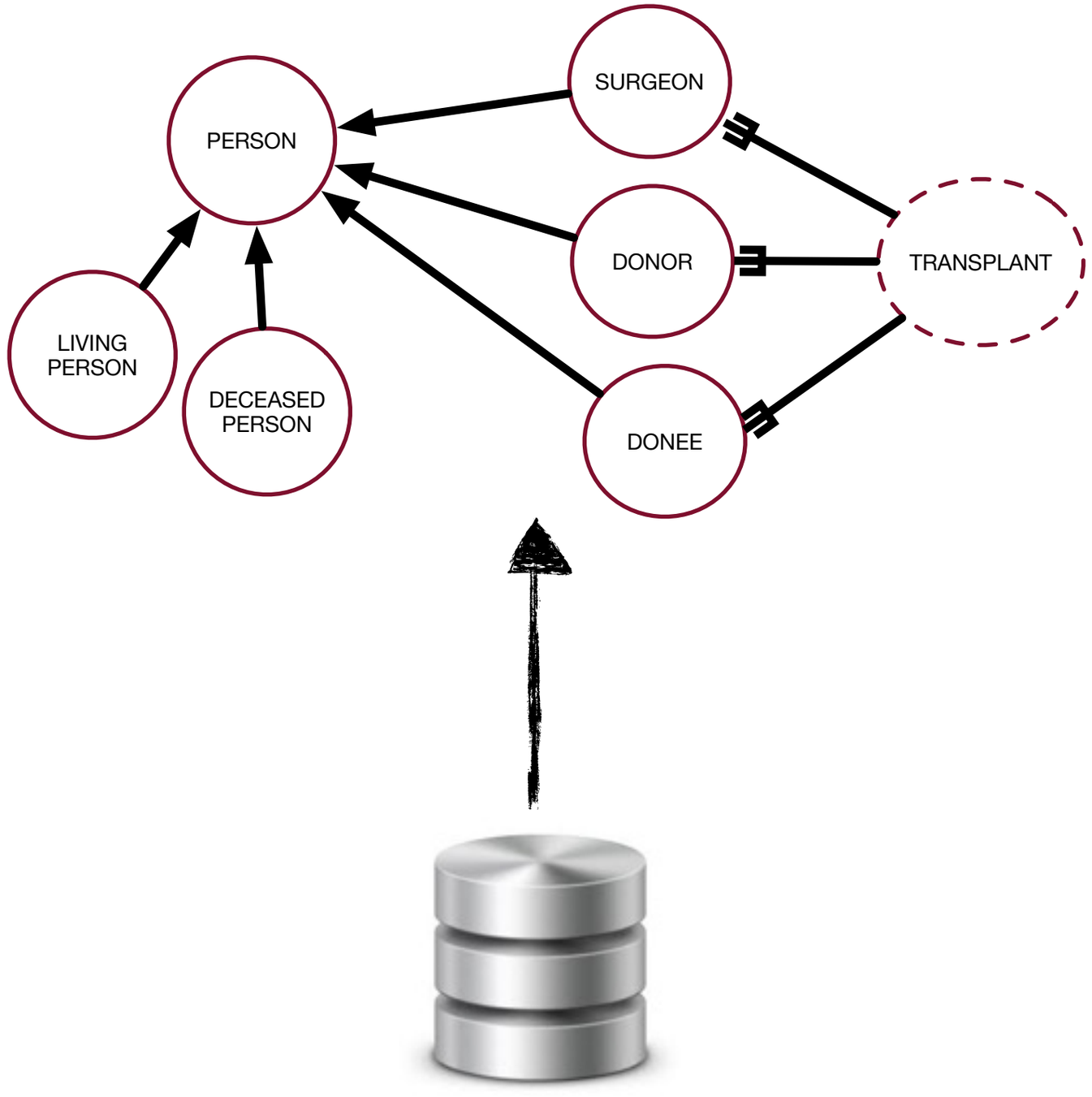
X

Real-World

(or **Ontological**)

Semantics


Isolatorweg
50 51
IJ
52
Noord
Noorderpark
Sloterdijk
De Vlugtlaan
Jan van Galenstraat
Postjesweg
Lelylaan
Heemstedestraat
Henk Sneevlietweg
Amstelveenseweg
Singelgracht
Prinsengracht
Keizersgracht
Herengracht
Singel
51 53 54
Centraal Station
Nieuwmarkt
Rokin
Waterlooplein
Weesperplein
Vijzelgracht
Wibautstraat
De Pijp
Amstel
Amstel
Europaplein
52
Zuid
RAI
Overamstel
Spaklerweg
IJmeer
Veensepolder
Diemen Zuid
Verrijn Stuartweg
Van der Madeweg
Duivendrecht
Ganzenhoef
Kraaiennest
Strandvliet
Gaasperplas
Bijlmer ArenA
Bullewijk
53
Holendrecht
50 54
Reigersbos
Gein


PERSON
SURGEON
DONOR
TRANSPLANT
LIVING PERSON
DECEASED PERSON
DONEE

Another look at data

by GEORGE H. MEALY
Computer Consultant
Scituate, Massachusetts

INTRODUCTION

particular ontology, we can avoid a quarrel by adopt-

*"data are fragments of a **theory of the real world**, and data processing juggles **representations** of these fragments of theory..."*

them in a somewhat new form may prove to be at least suggestive.

To begin on a philosophical plane, let us note that we usually behave as if there were three realms of interest in data processing: the real world itself, ideas about it existing in the minds of men, and symbols on paper or some other storage medium. The lat-

Toward a theory of data

Relations

To fix our ideas, consider the following example of genealogical data, taken from Reference 2:

Another look at data

by GEORGE H. MEALY
Computer Consultant
Scituate, Massachusetts

INTRODUCTION

particular ontology, we can avoid a quarrel by adopt-

*"data are fragments of a theory of the real world, and data processing juggles representations of these fragments of theory...**The issue is ontology, or the question of what exists.**"*

them in a somewhat new form may prove to be at least suggestive.

To begin on a philosophical plane, let us note that we usually behave as if there were three realms of interest in data processing: the real world itself, ideas about it existing in the minds of men, and symbols on paper or some other storage medium. The lat-

Toward a theory of data

Relations

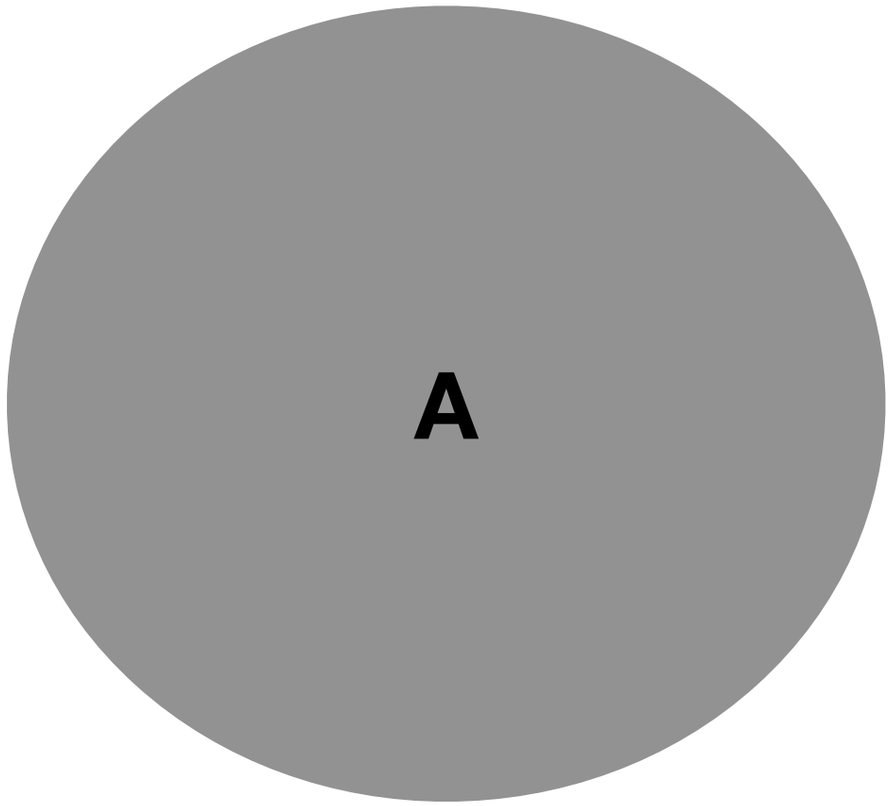
To fix our ideas, consider the following example of genealogical data, taken from Reference 2:

ontology ≈

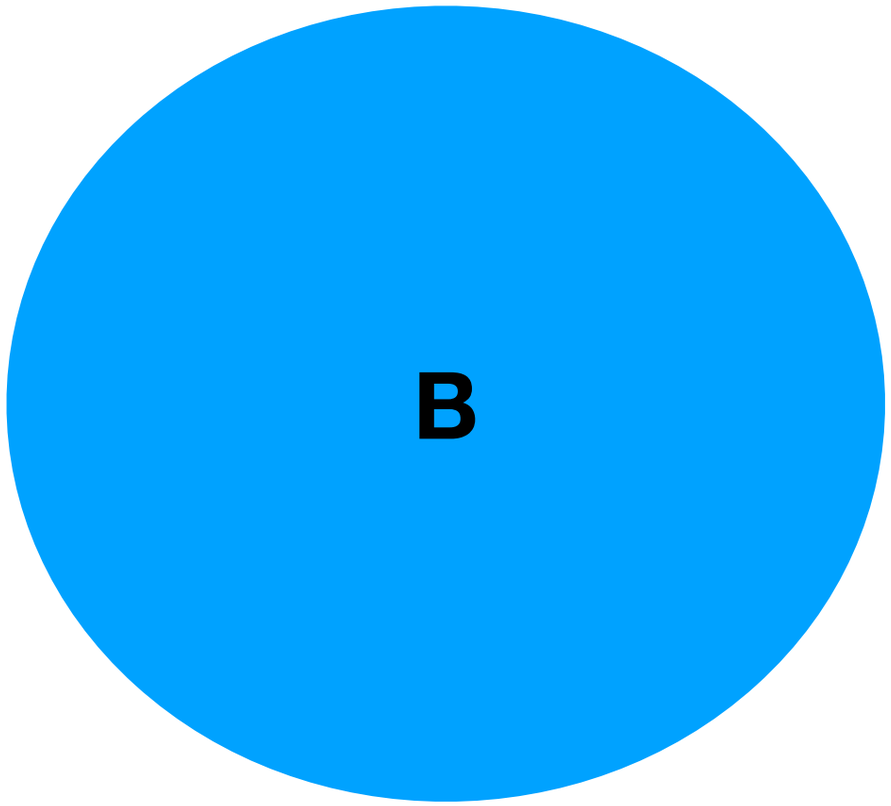
A theory about the kinds of entities and their ties that are assumed to exist by a given description of reality

ontology ≈

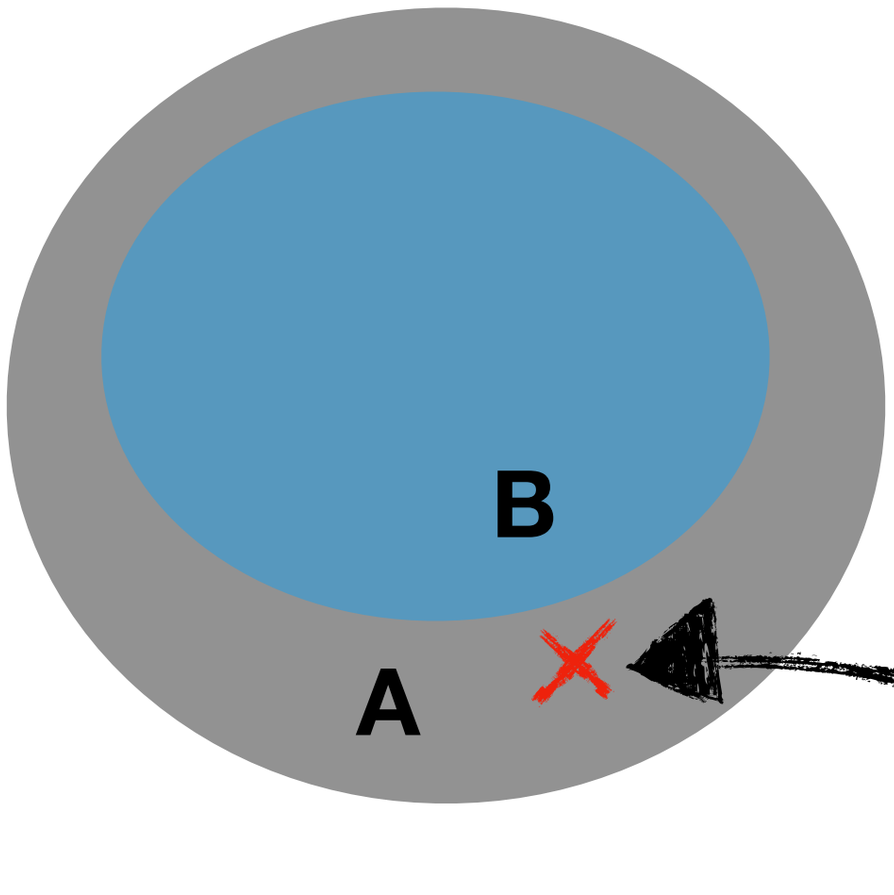
~~A theory about the kinds of entities and their ties that are assumed to exist by~~ a given description of reality



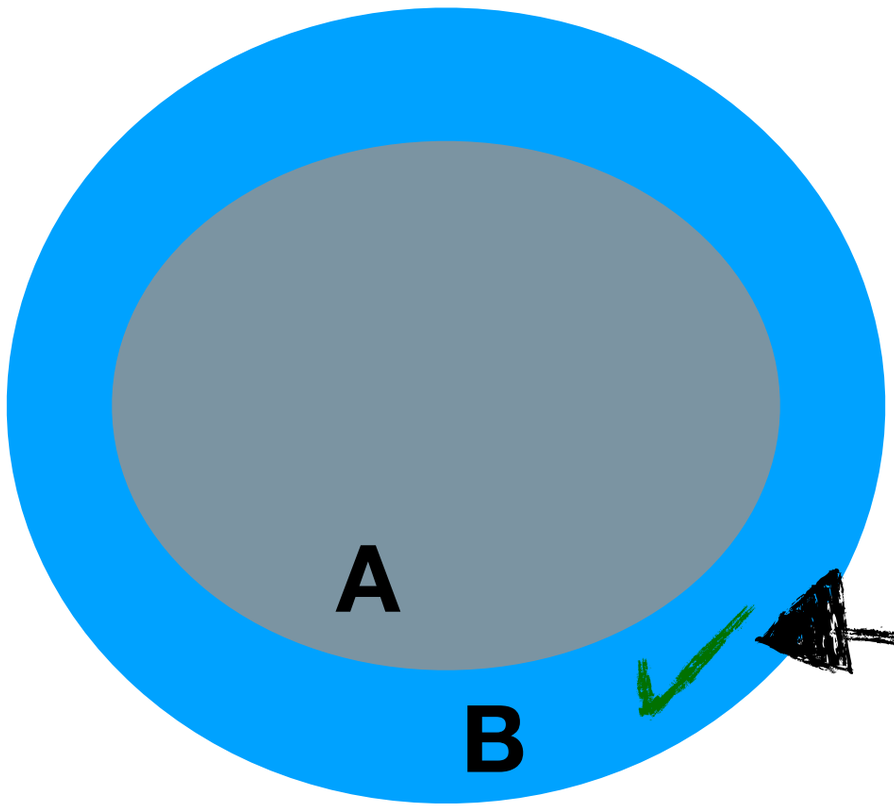
Possible Interpretations of a Model



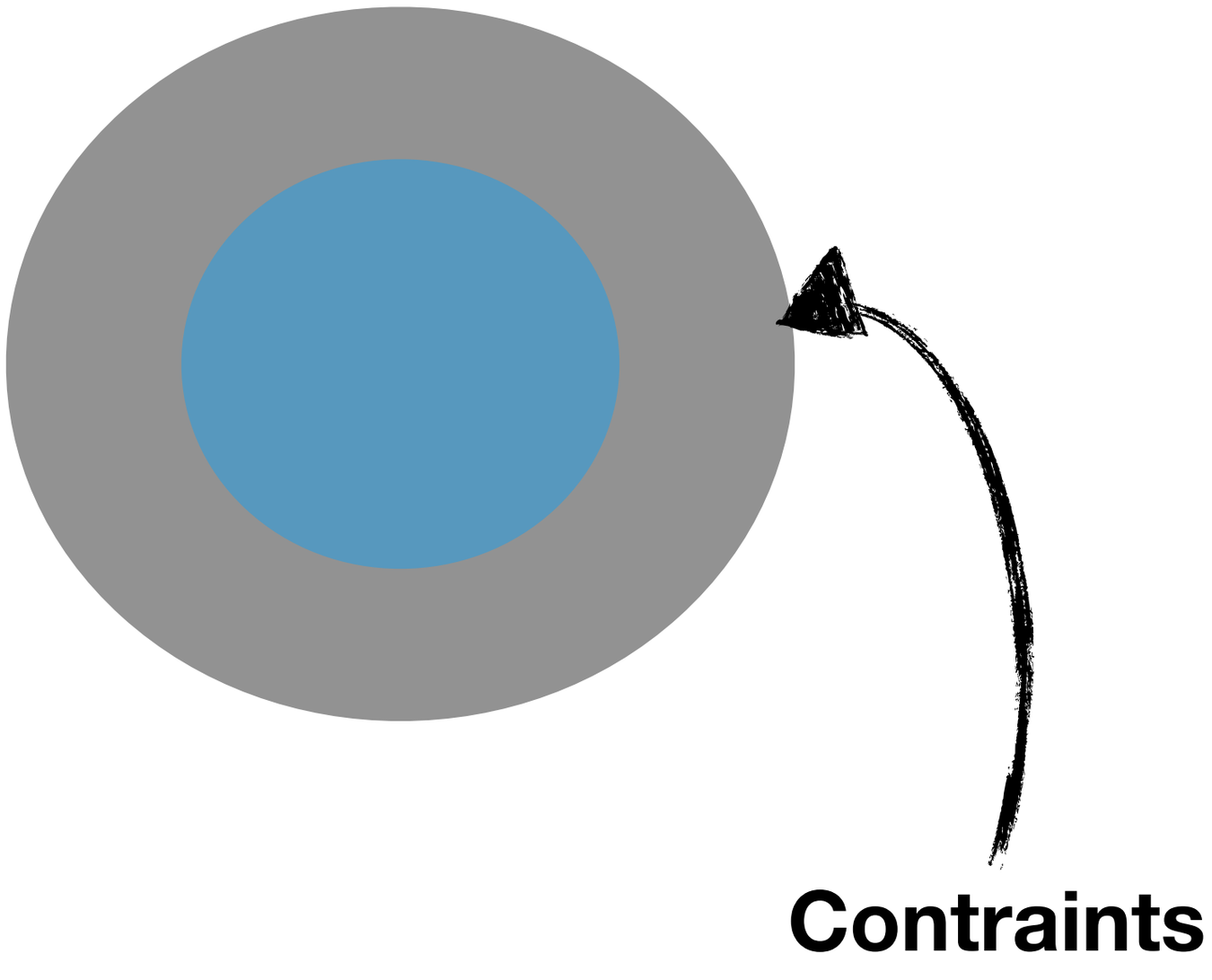
Intended Interpretations of that Model



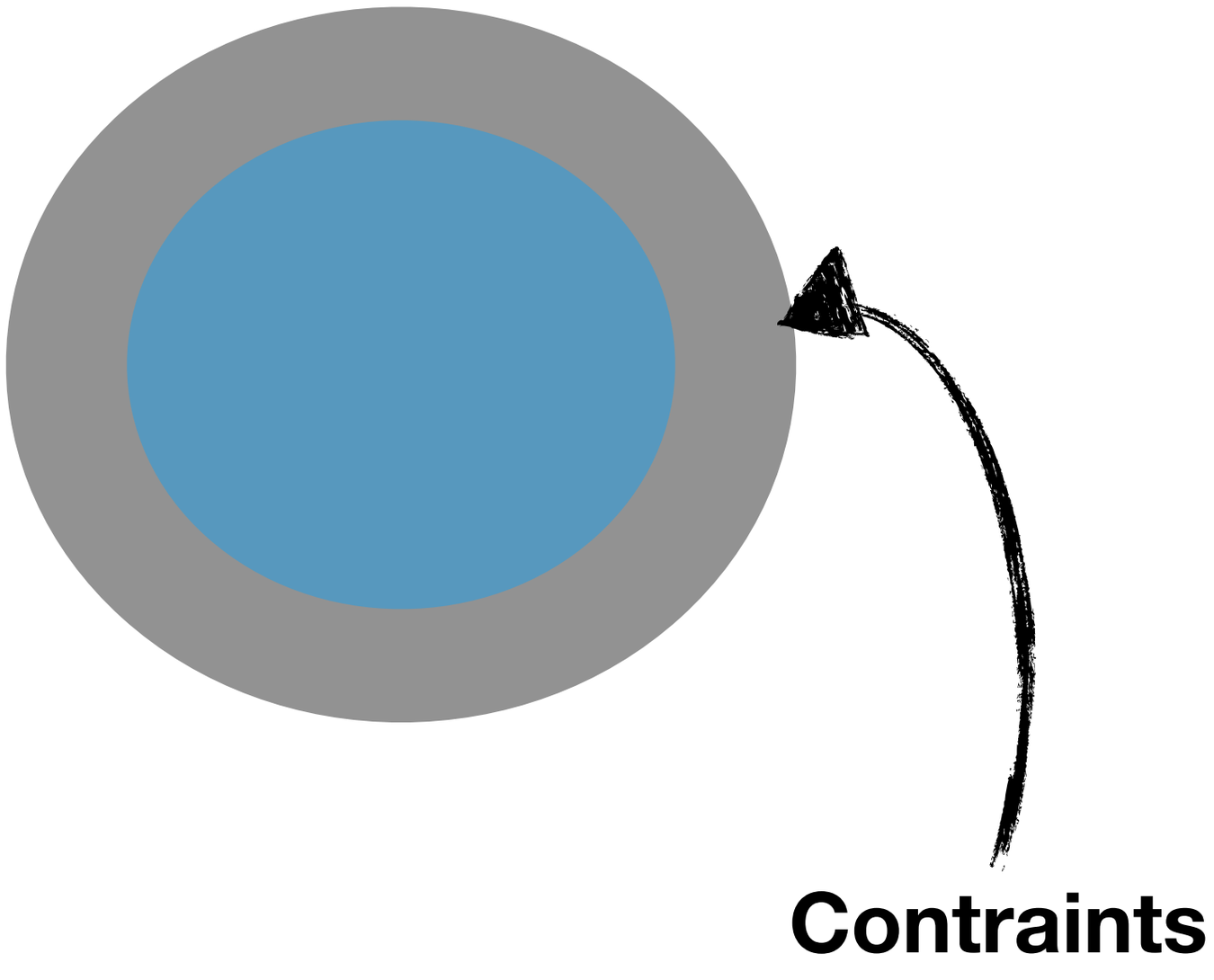
Under-constrained Model

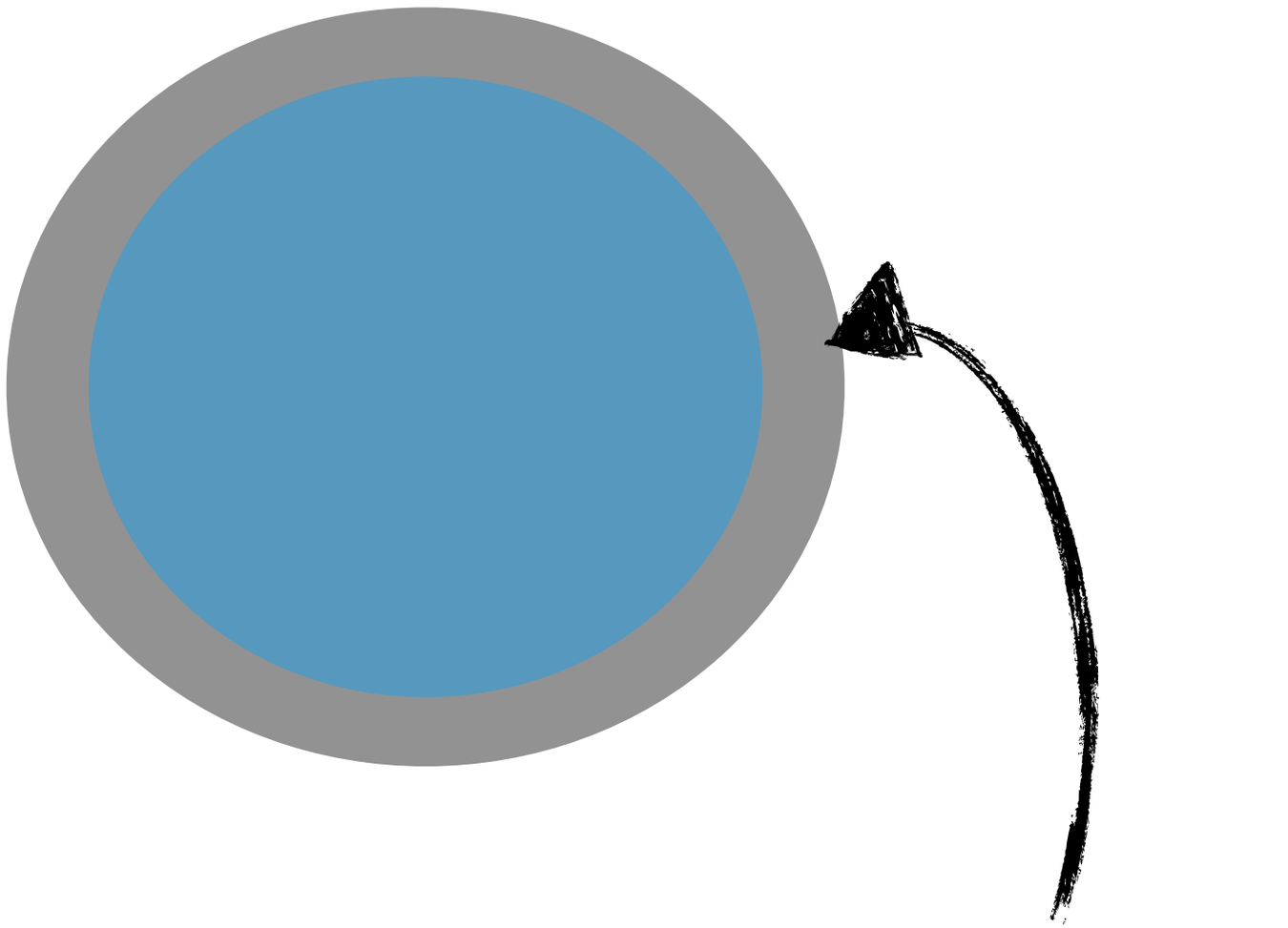


Over-constrained Model

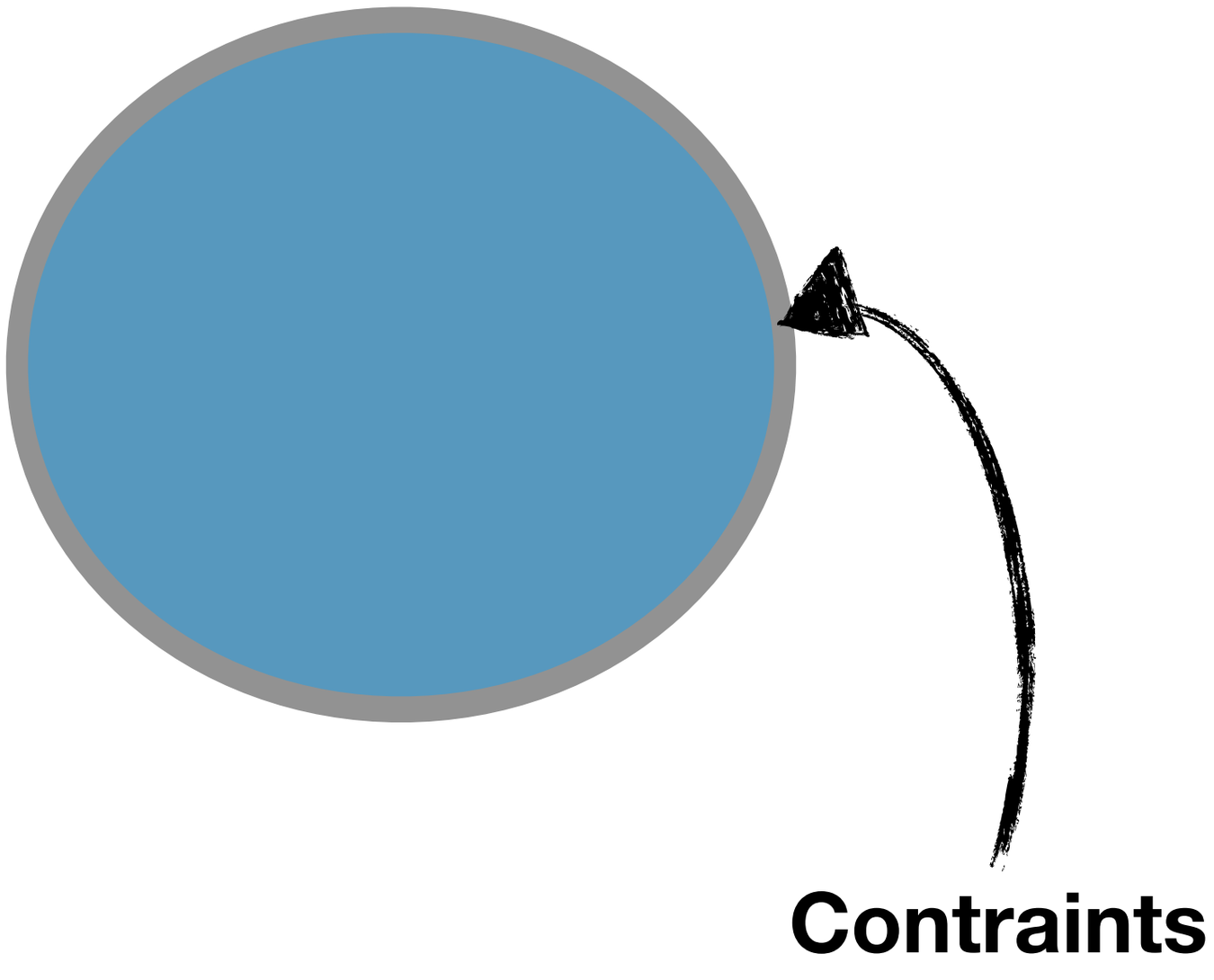


Contraints

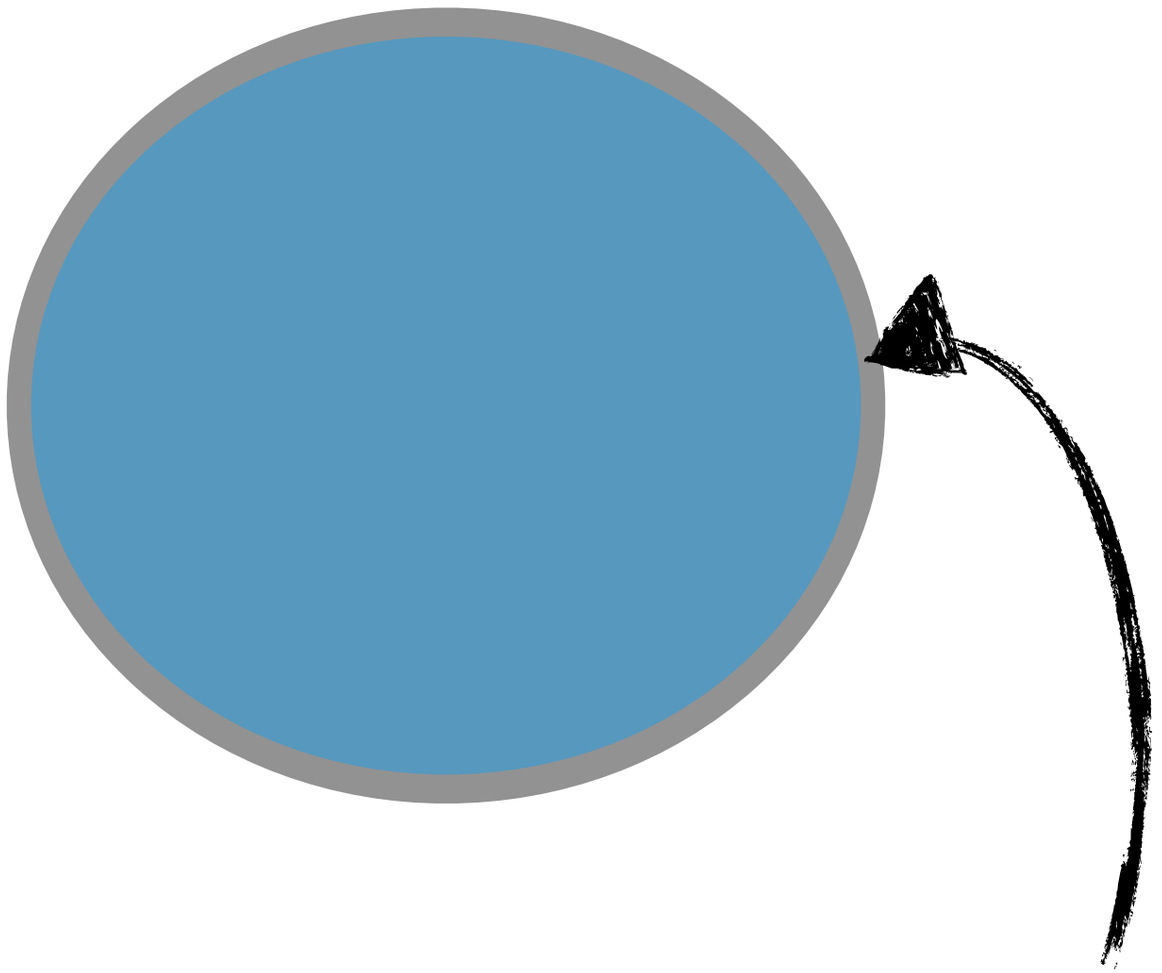

Contraints



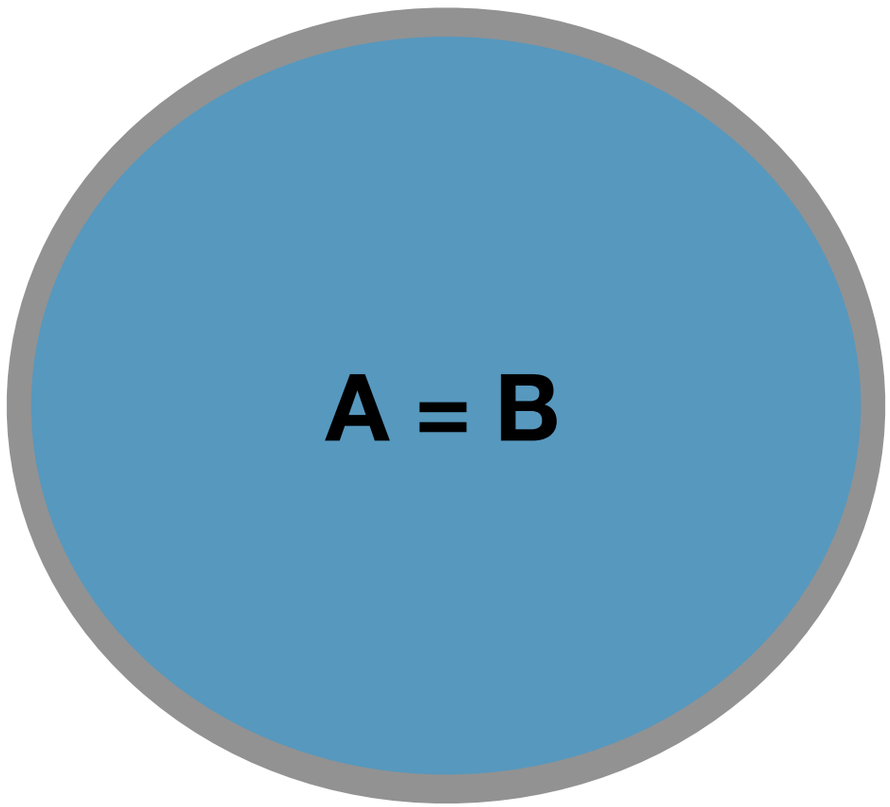
Contraints

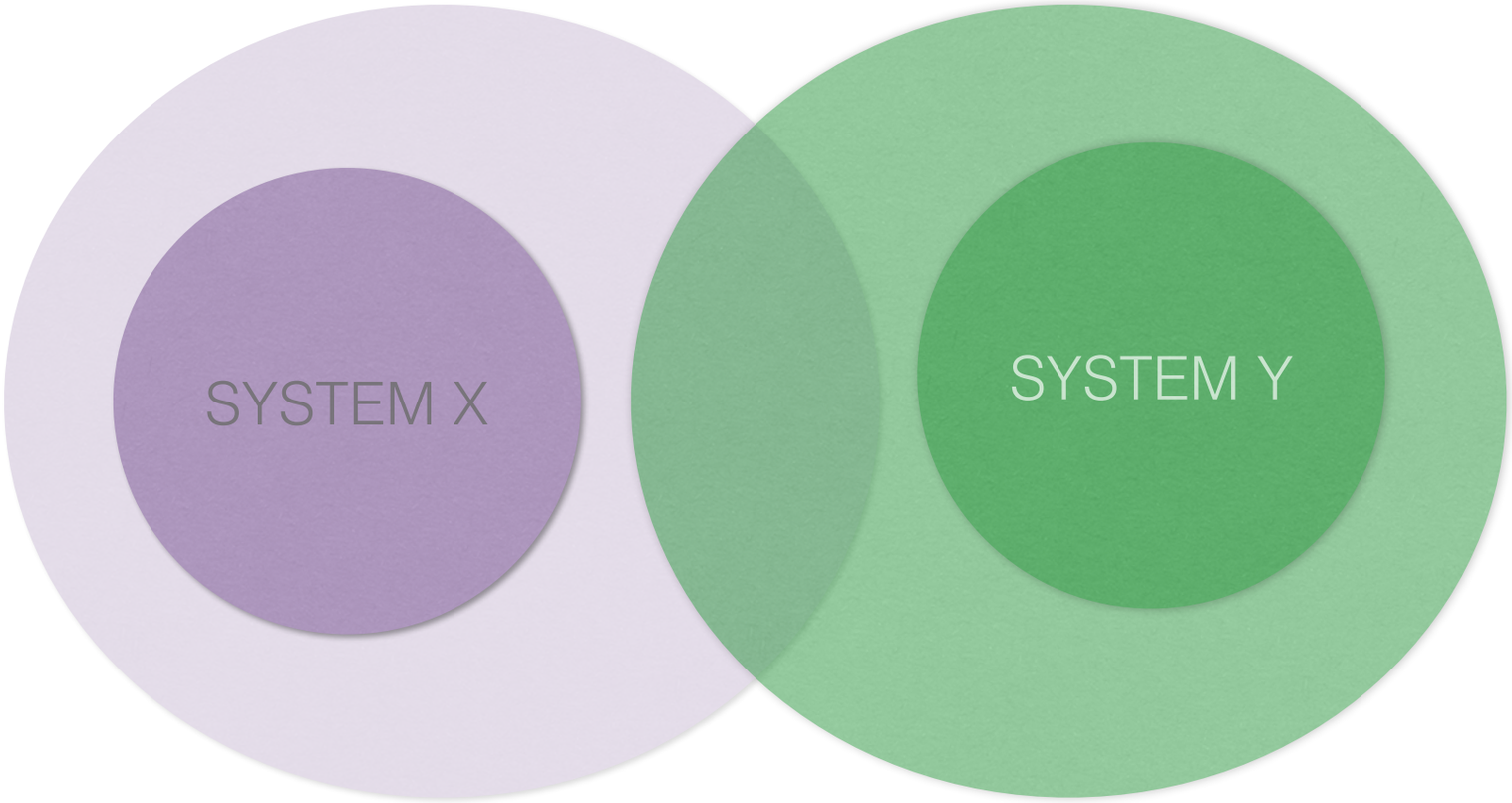


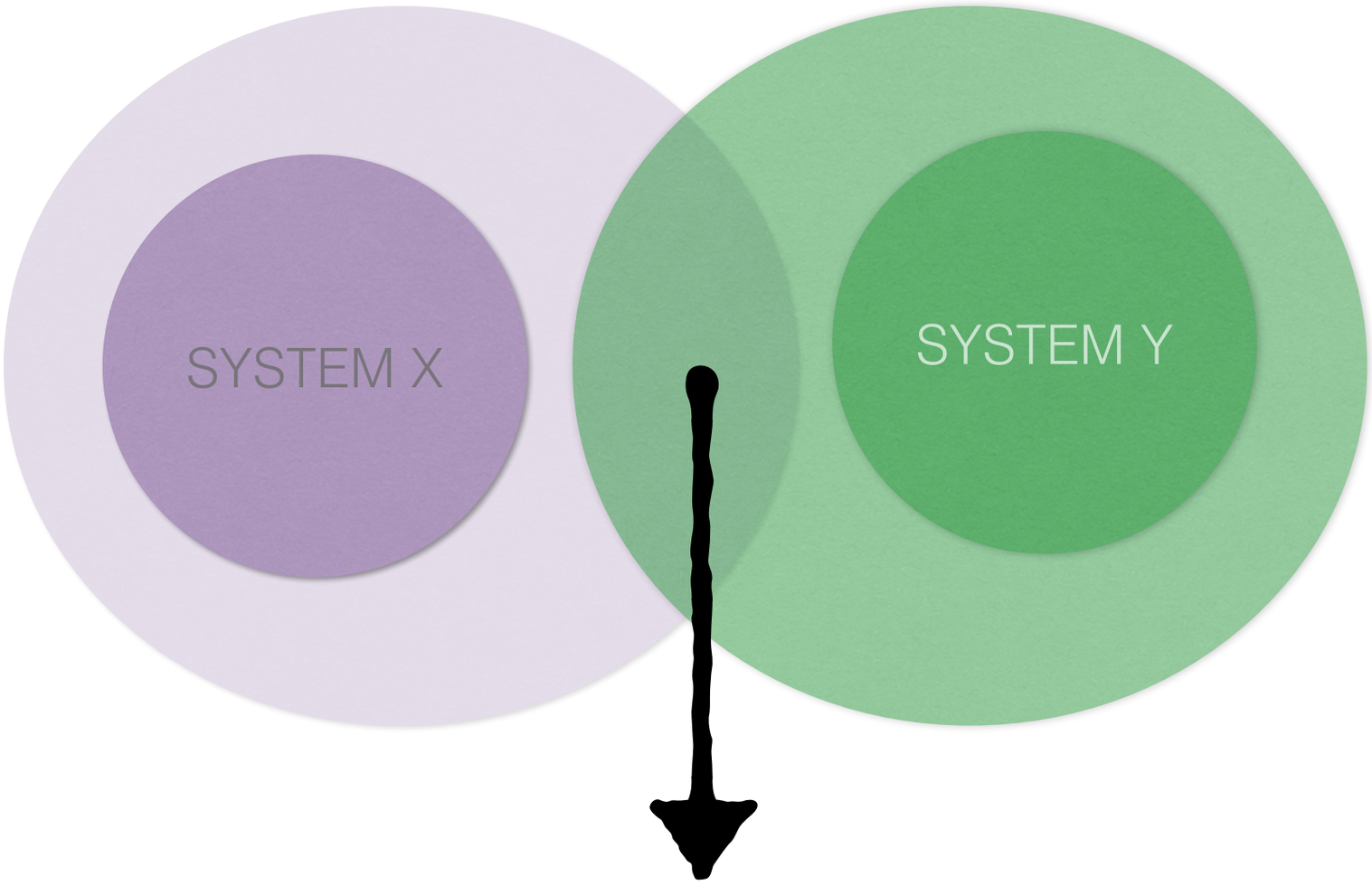
Contraints



Ontological Commitment


A = B


SYSTEM X
SYSTEM Y

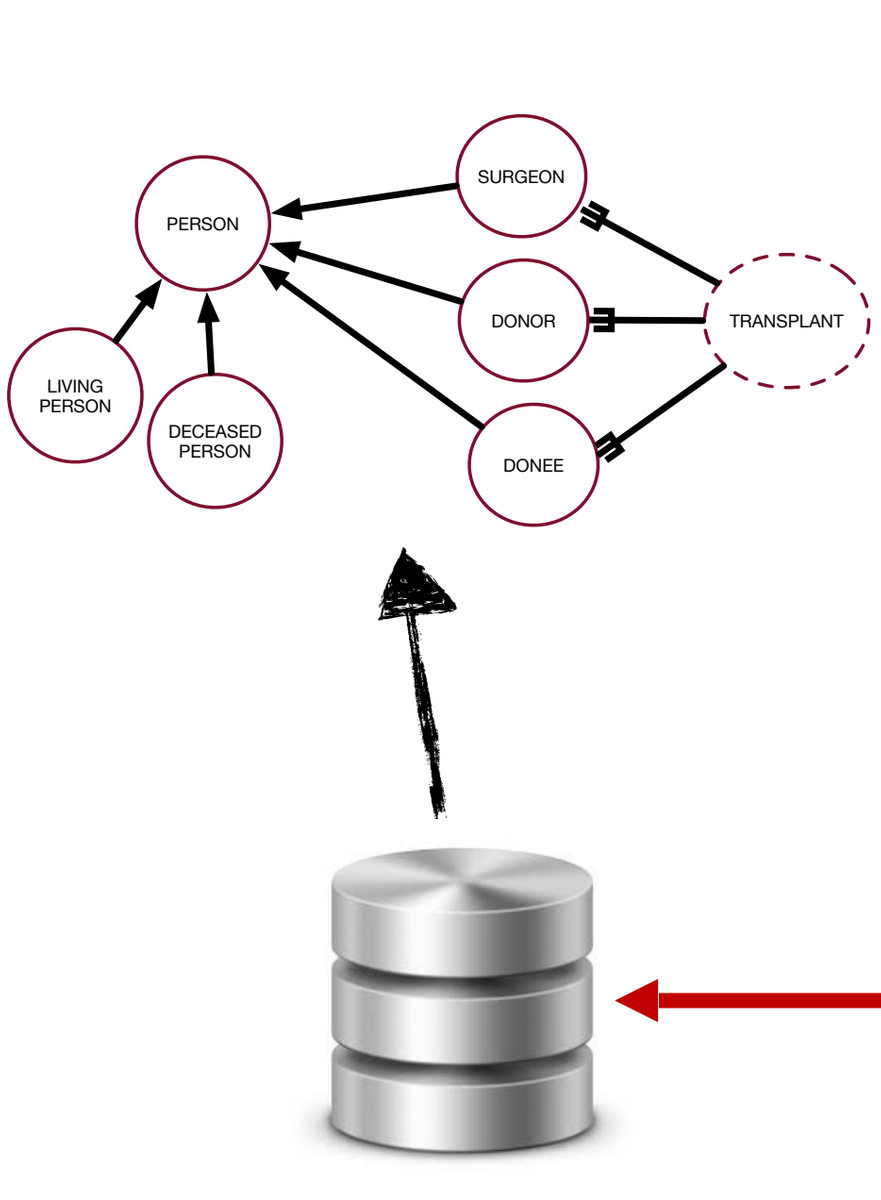


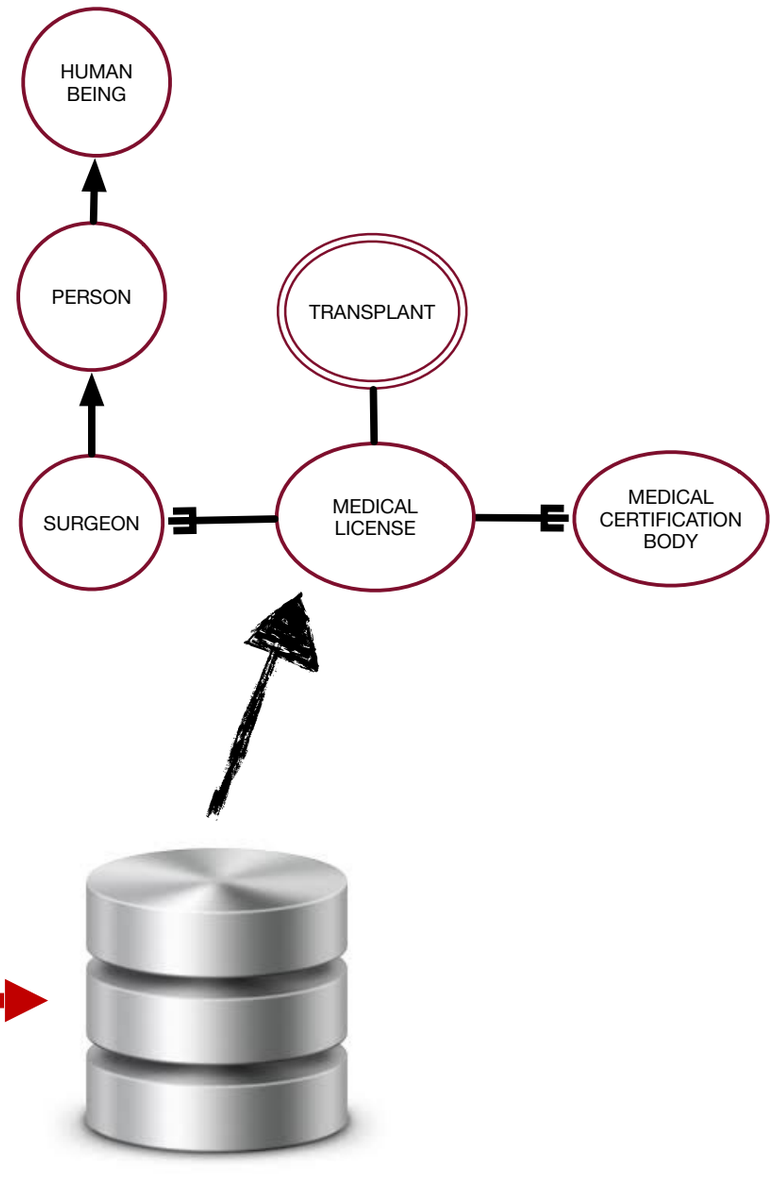
FALSE AGREEMENT

Semantic Interoperability

=

relating different
worldviews, i.e., different
ontologies


PERSON
SURGEON
DONOR
TRANSPLANT
LIVING PERSON
DECEASED PERSON
DONEE


HUMAN BEING
PERSON
TRANSPLANT
SURGEON
MEDICAL LICENSE
MEDICAL CERTIFICATION BODY

?

PERSON

SURGEON

DONOR

DONEE

TRANSPLANT

LIVING PERSON

DECEASED PERSON

HUMAN BEING

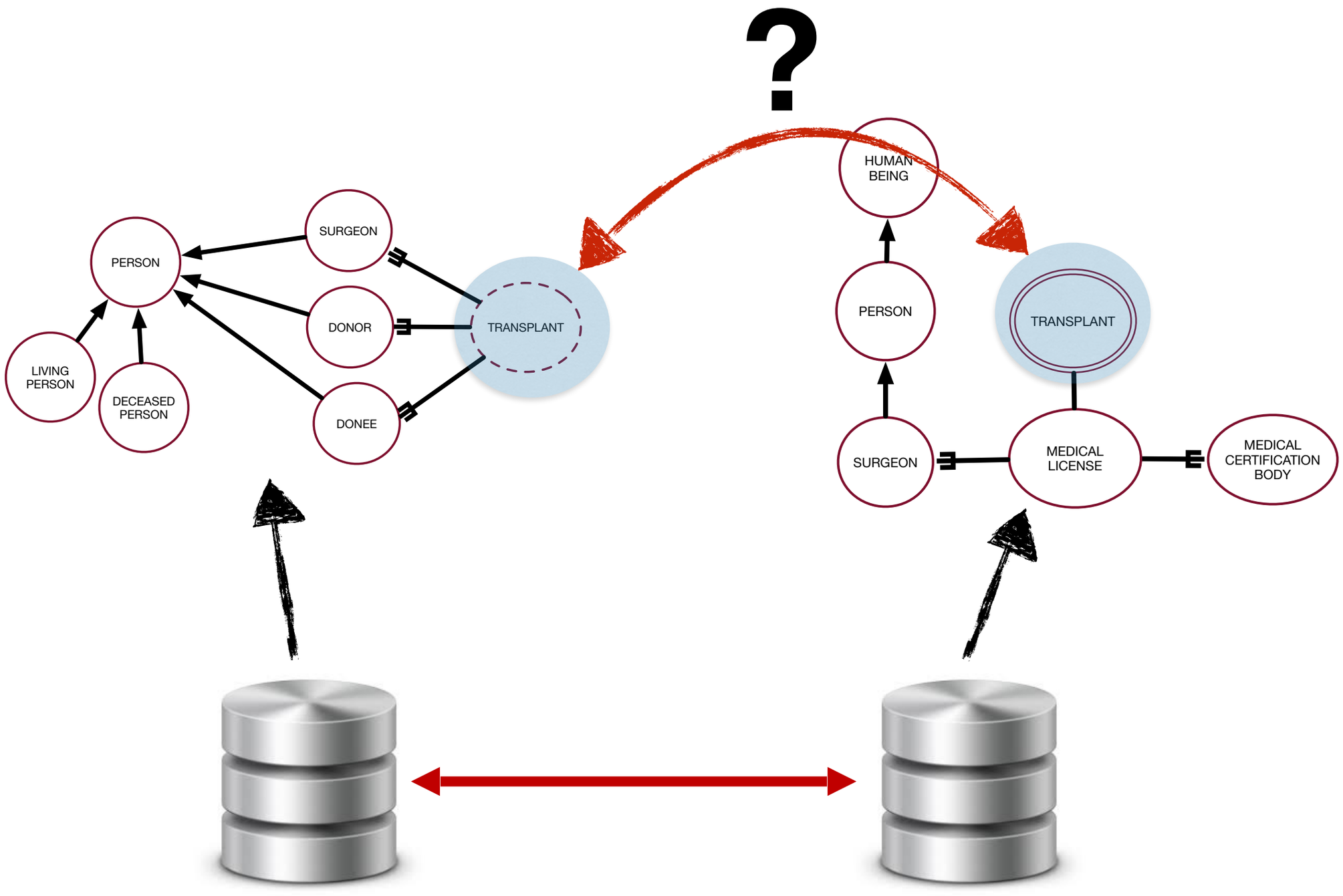
PERSON

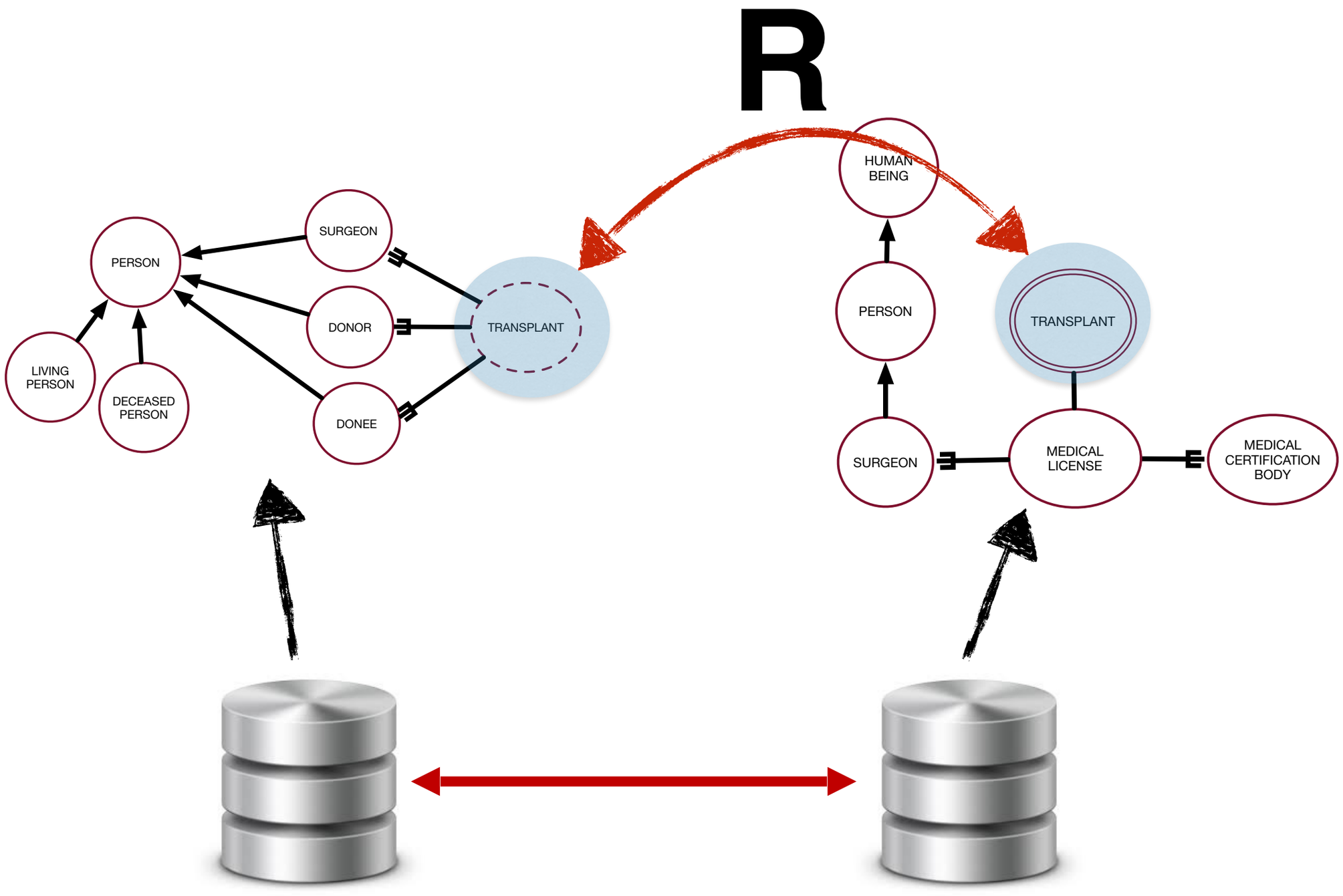
TRANSPLANT

SURGEON

MEDICAL LICENSE

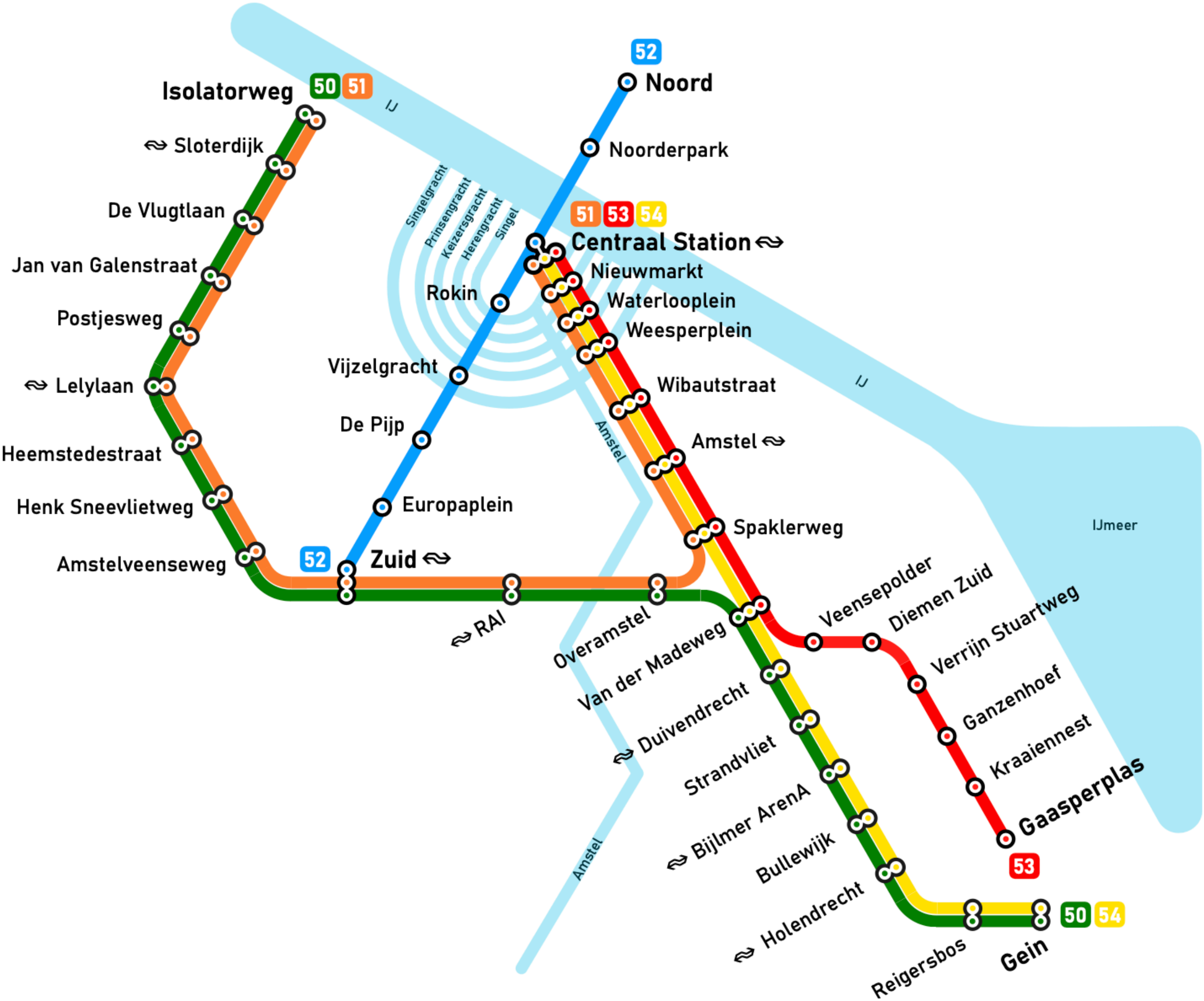
MEDICAL CERTIFICATION BODY

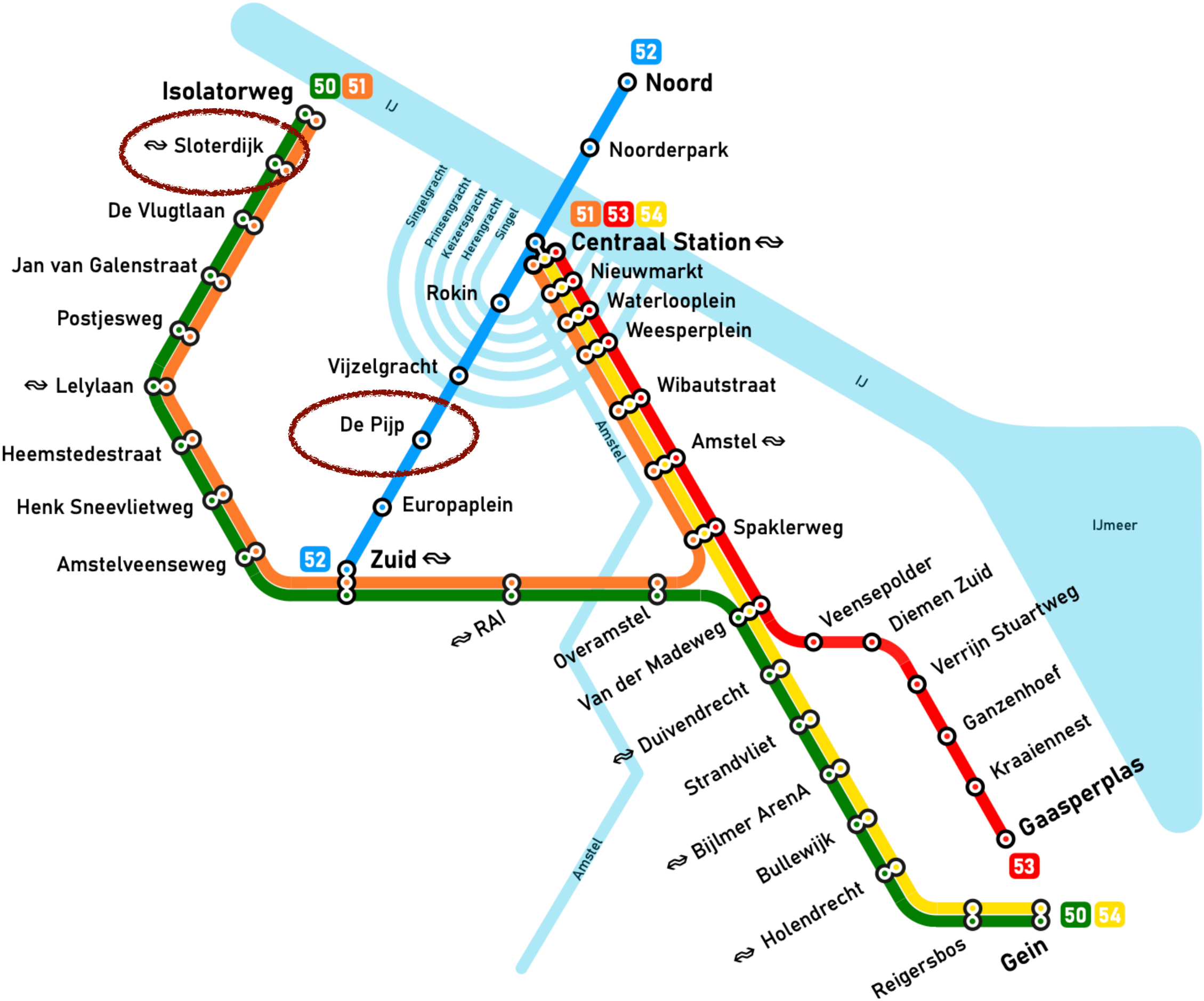

?
HUMAN BEING
PERSON
SURGEON
DONOR
DONEE
TRANSPLANT
LIVING PERSON
DECEASED PERSON
PERSON
TRANSPLANT
SURGEON
MEDICAL LICENSE
MEDICAL CERTIFICATION BODY


R
HUMAN BEING
PERSON
SURGEON
DONOR
DONEE
LIVING PERSON
DECEASED PERSON
TRANSPLANT
PERSON
TRANSPLANT
SURGEON
MEDICAL LICENSE
MEDICAL CERTIFICATION BODY

Ontology ≈

An area devoted to developing these domain-independent “toolboxes” with “tools”for supporting ontological analysis


Isolatorweg
50 51
IJ
52
Noord
Noorderpark
Sloterdijk
De Vlugtlaan
Jan van Galenstraat
Postjesweg
Lelylaan
Heemstedestraat
Henk Sneevlietweg
Amstelveenseweg
Singelgracht
Prinsengracht
Keizersgracht
Herengracht
Singel
51 53 54
Centraal Station
Nieuwmarkt
Rokin
Waterlooplein
Weesperplein
Vijzelgracht
Wibautstraat
De Pijp
Amstel
Amstel
Europaplein
52
Zuid
RAI
Overamstel
Spaklerweg
IJmeer
Veensepolder
Diemen Zuid
Verrijn Stuartweg
Van der Madeweg
Duivendrecht
Ganzenhoef
Kraaiennest
Strandvliet
Gaasperplas
Bijlmer ArenA
Bullewijk
53
Holendrecht
50 54
Reigersbos
Gein


Isolatorweg
50 51
IJ
52
Noord
Noorderpark
Sloterdijk
De Vlugtlaan
Jan van Galenstraat
Postjesweg
Lelylaan
Heemstedestraat
Henk Sneevlietweg
Amstelveenseweg
Singelgracht
Prinsengracht
Keizersgracht
Herengracht
Singel
51 53 54
Centraal Station
Nieuwmarkt
Waterlooplein
Weesperplein
Rokin
Vijzelgracht
De Pijp
Europaplein
52
Zuid
RAI
Overamstel
Amstel
Wibautstraat
Amstel
Spaklerweg
IJmeer
Veensepolder
Diemen Zuid
Verrijn Stuartweg
Ganzenhoef
Kraaiennest
Gaasperplas
53
Van der Madeweg
Duivendrecht
Strandvliet
Bijlmer ArenA
Bullewijk
Holendrecht
Reigersbos
Gein
50 54

“The **ontological** approach to **explanation**” by

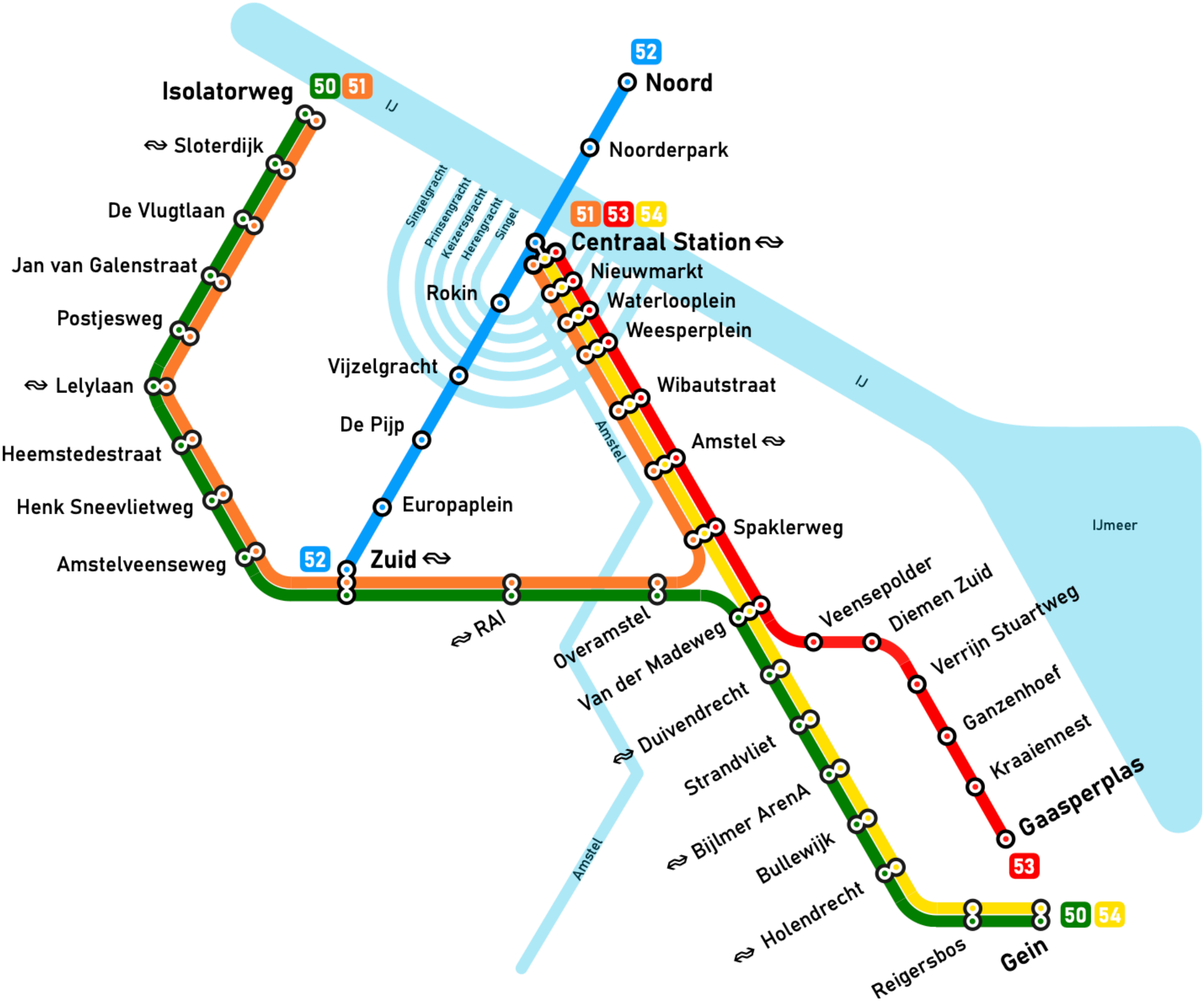
T.Y. Cao(2004). Ontology and scientific explanation. Oxford University Press

“the notion of a basic ontology in a scientific theory refers to the irreducible conceptual element…of what is assumed to…exist in the domain under investigation. As a representation of deep reality, the basic ontology enjoys a great explanatory power. That is, all appearances should be derivable from it as a result of its behavior”.

“whenever we have something important but difficult to understand, we should focus our attention on finding what the primary entities in the domain under investigation are…”

“…Discovering these entities and their intrinsic and structural properties…is the real work of science…”

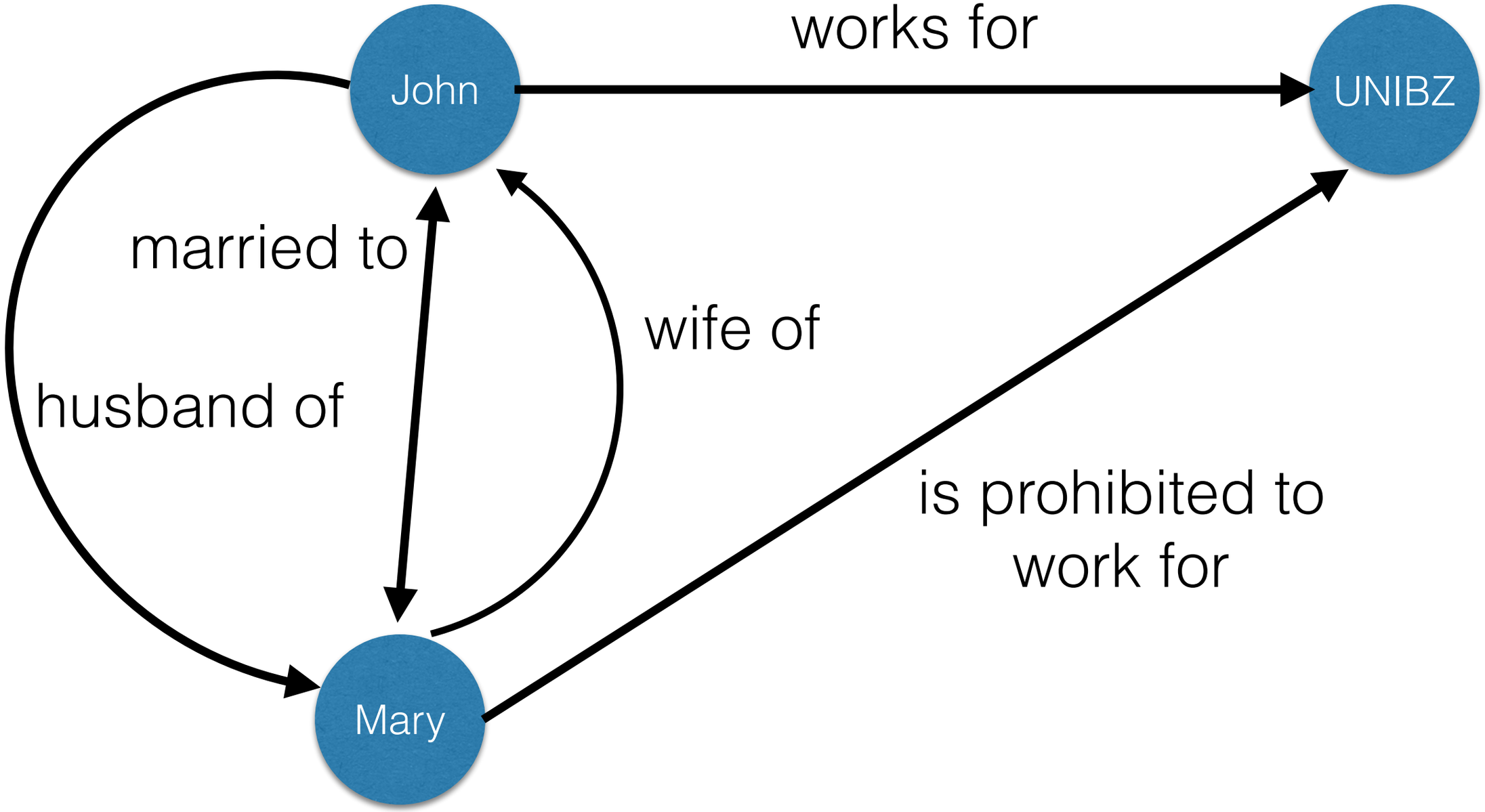
“Mathematical formalisms and universal laws and principles are relevant and important only when they have a firm ontological basis.”

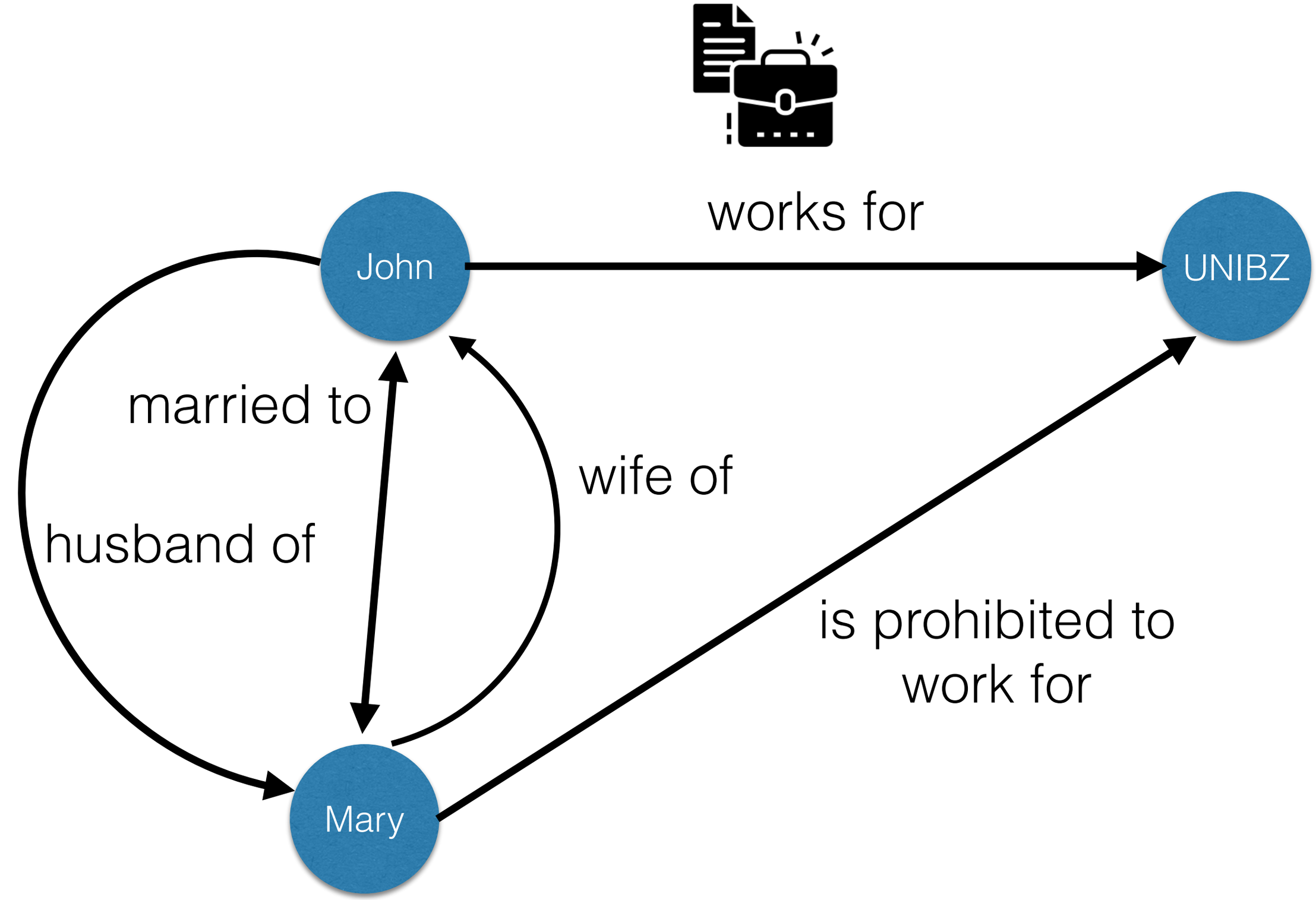

Isolatorweg
50 51
IJ
52
Noord
Noorderpark
Sloterdijk
De Vlugtlaan
Jan van Galenstraat
Postjesweg
Lelylaan
Heemstedestraat
Henk Sneevlietweg
Amstelveenseweg
Singelgracht
Prinsengracht
Keizersgracht
Herengracht
Singel
51 53 54
Centraal Station
Nieuwmarkt
Rokin
Waterlooplein
Weesperplein
Vijzelgracht
Wibautstraat
De Pijp
Amstel
Amstel
Europaplein
52
Zuid
RAI
Overamstel
Spaklerweg
IJmeer
Veensepolder
Diemen Zuid
Verrijn Stuartweg
Van der Madeweg
Duivendrecht
Ganzenhoef
Kraaiennest
Strandvliet
Gaasperplas
Bijlmer ArenA
Bullewijk
53
Holendrecht
50 54
Reigersbos
Gein

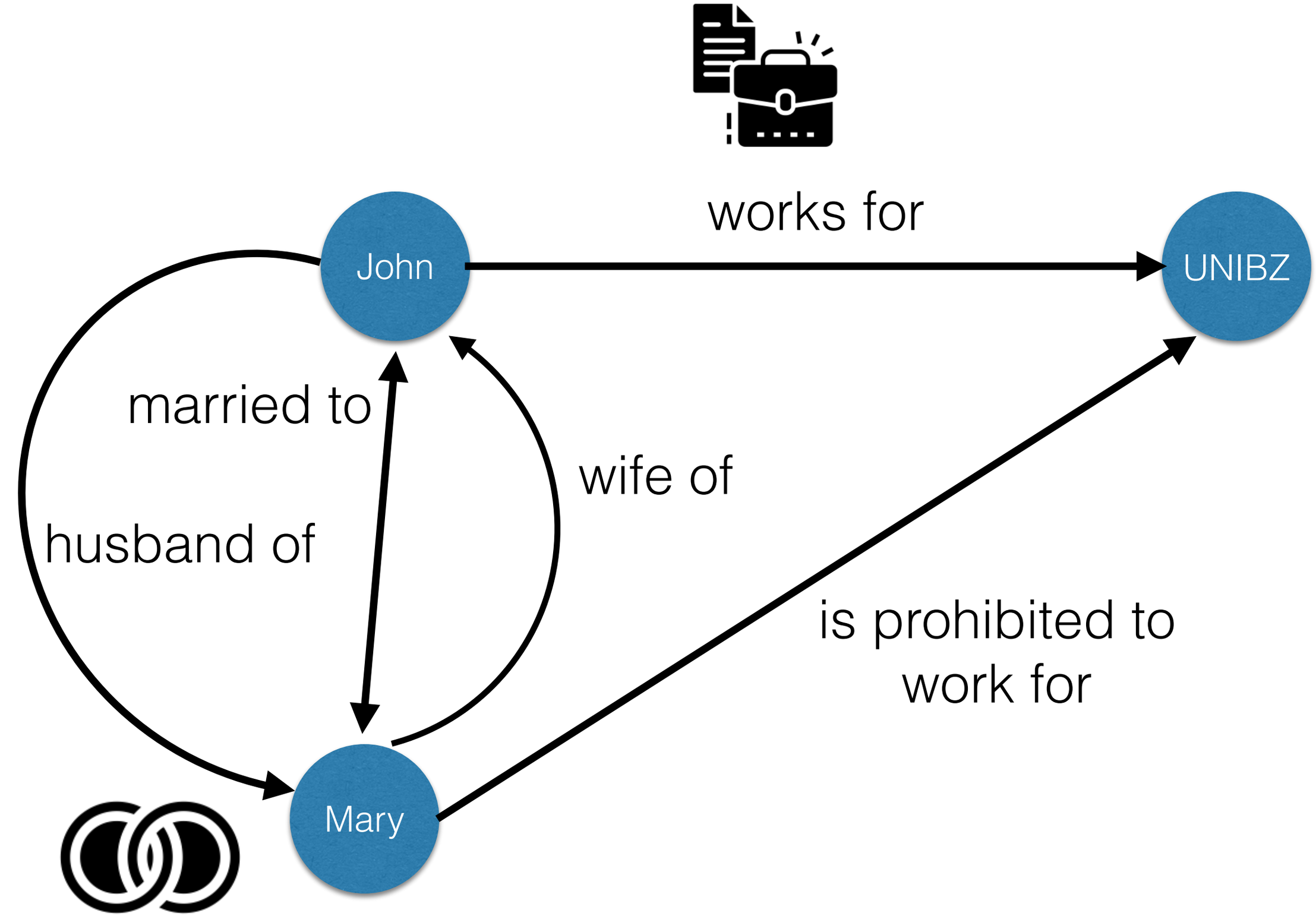
Truth-bearers

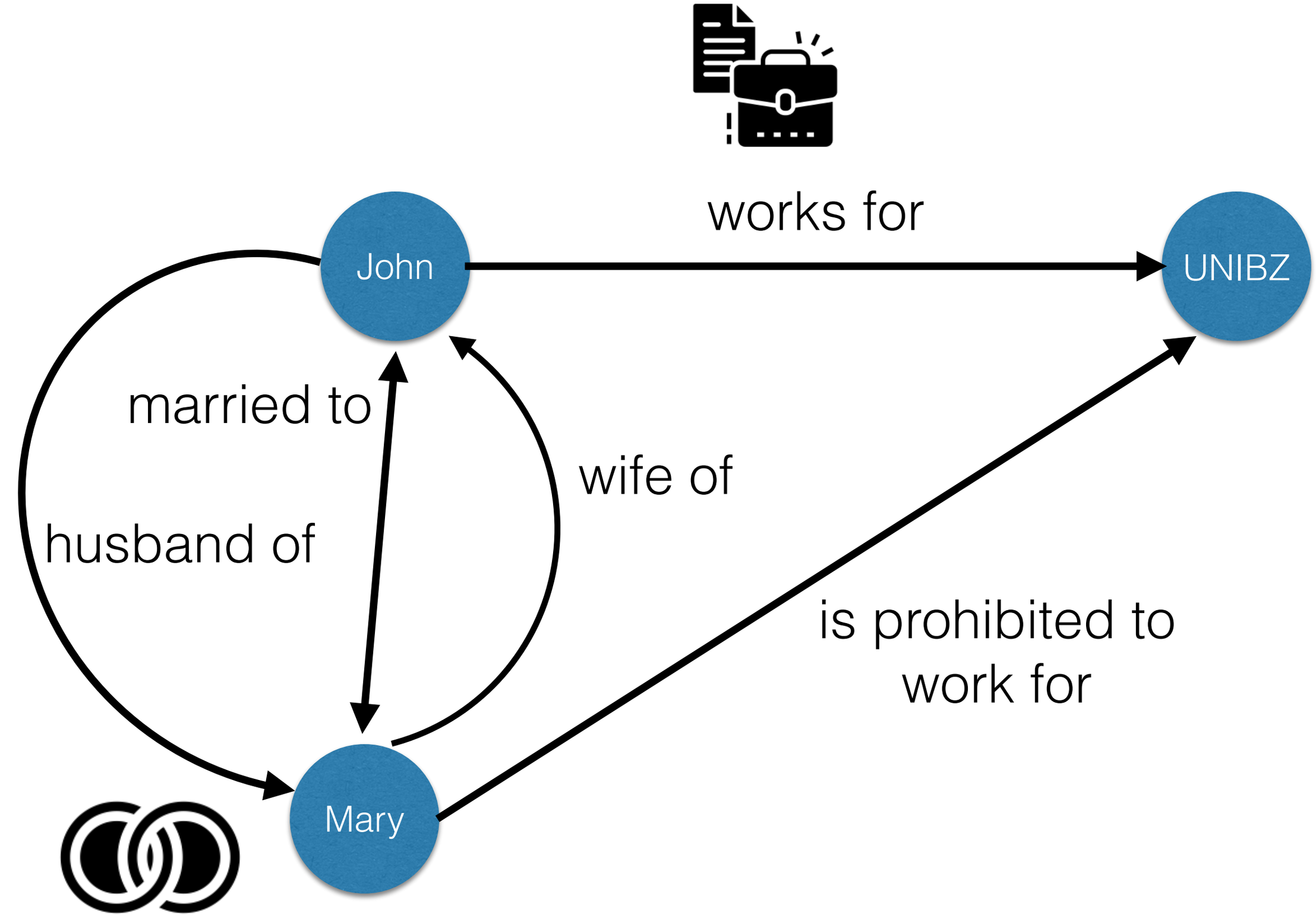
X

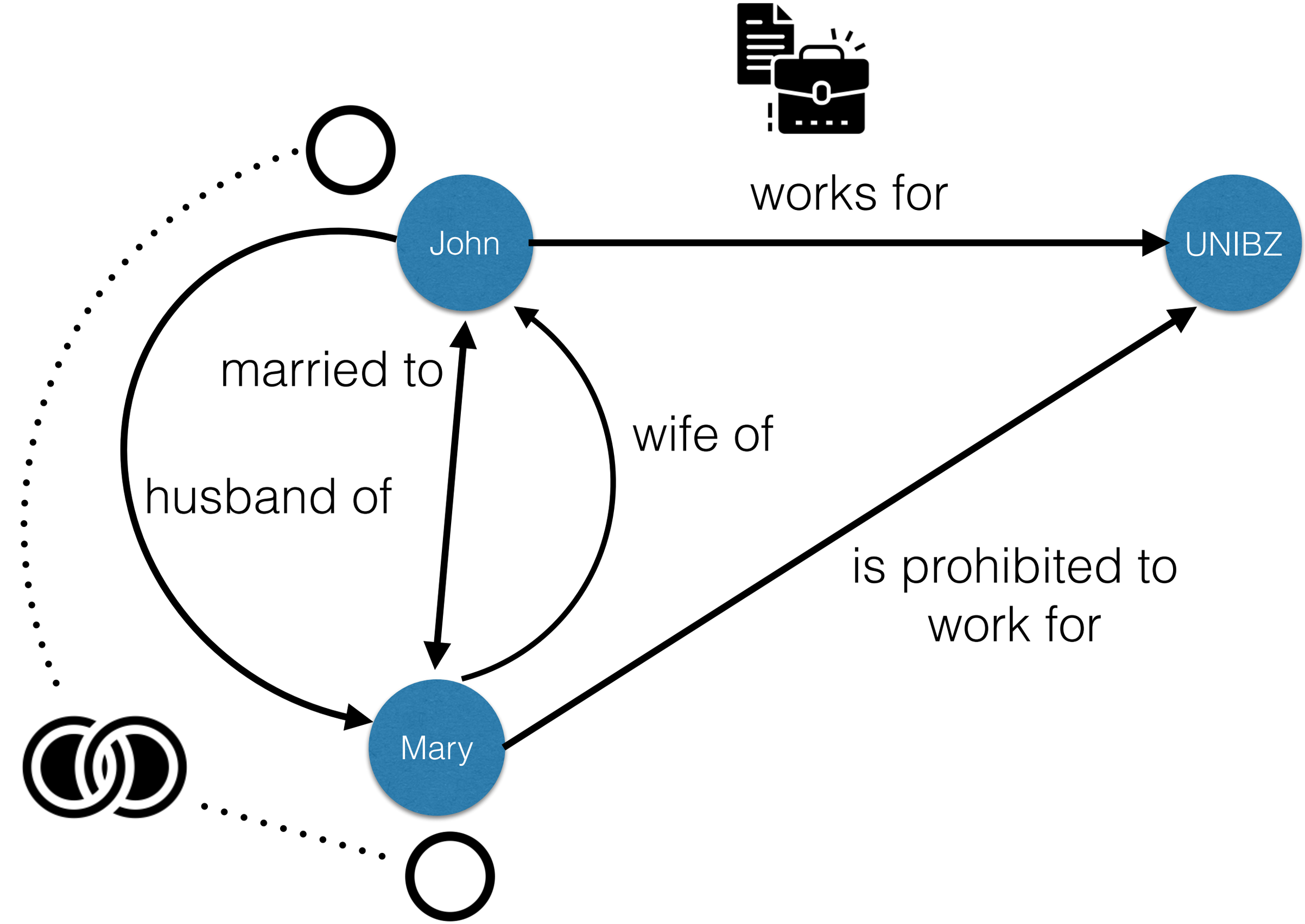
Truth-makers

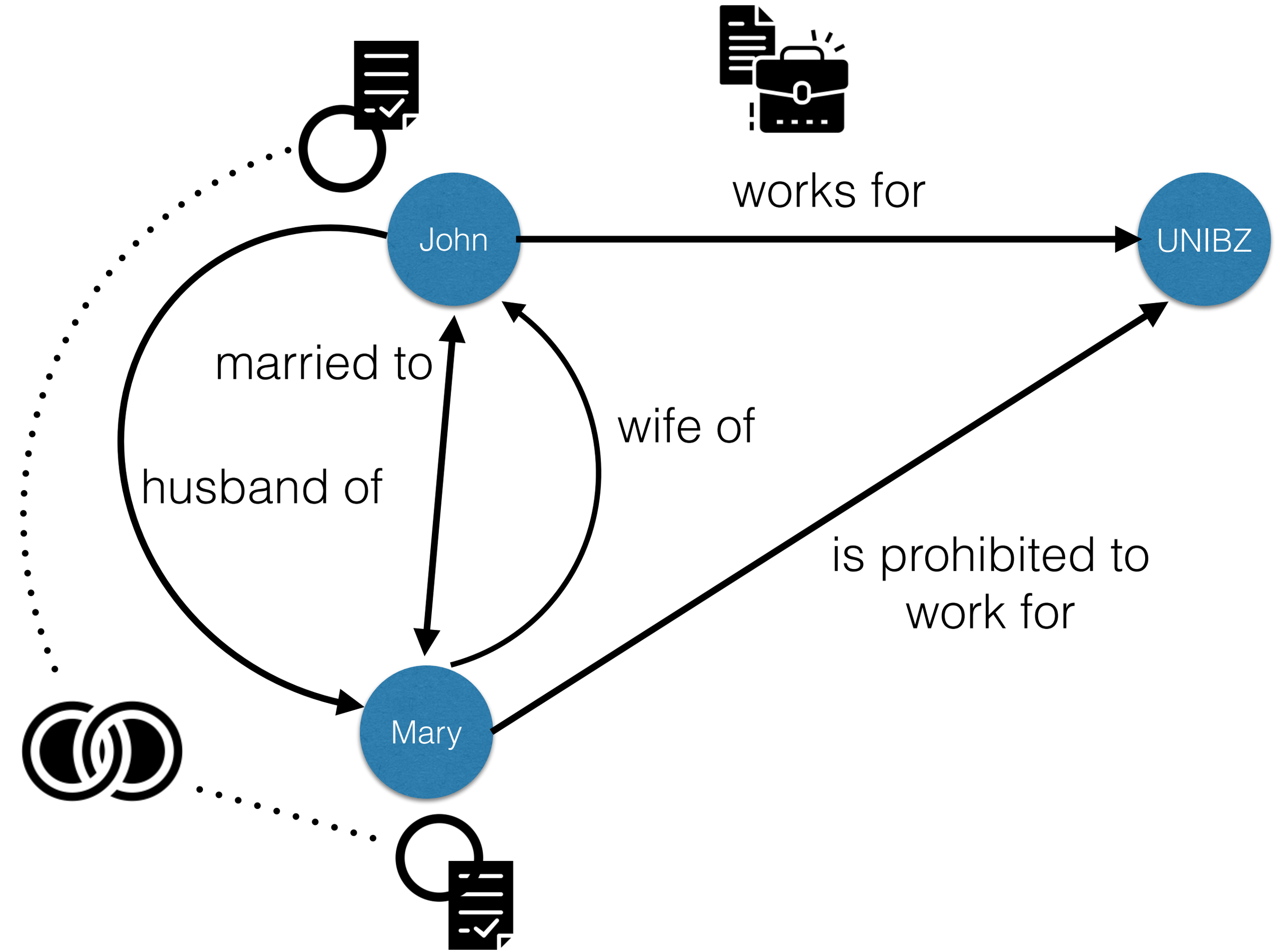

works for
John
UNIBZ
married to
husband of
wife of
is prohibited to
work for
Mary


works for
John
UNIBZ
married to
husband of
wife of
is prohibited to
work for
Mary


works for
John
UNIBZ
married to
husband of
wife of
is prohibited to
work for
Mary


works for
John
UNIBZ
married to
husband of
wife of
is prohibited to
work for
Mary


works for
John
UNIBZ
married to
husband of
wife of
is prohibited to
work for
Mary


works for
John
UNIBZ
married to
husband of
wife of
is prohibited to
work for
Mary

Structuring Function
(of a Conceptual Model)
X
Ontological Function
(of a Conceptual Model)

Truth-bearers
(Descriptive)

X

Truth-makers
(Explanatory)

RESEARCH ARTICLE

Semantic Interoperability: Ontological Unpacking of a Viral Conceptual Model

Anna Bernasconi[1,4*], Giancarlo Guizzardi[2,3], Oscar Pastor[4] and Veda C. Storey[5]

*Correspondence:
anna.bernasconi@polimi.it
[1]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy
[4]PROS Research Center & VRAIN Research Institute, Universidad Politècnica de València, Valencia, Spain
Full list of author information is available at the end of the article

Abstract

Background: Genomics and virology are unquestionably important, but complex, domains being investigated by a large number of scientists. The need to facilitate and support work within these domains requires sharing of databases, although it is often difficult to do so because of the different ways in which data is represented across the databases. To foster semantic interoperability, models are needed that provide a deep understanding and interpretation of the concepts in a domain, so that the data can be consistently interpreted among researchers.

Results: In this research, we propose the use of conceptual models to support semantic interoperability among databases and assess their ontological clarity to support their effective use. This modeling effort is illustrated by its application to the Viral Conceptual Model (VCM) that captures and represents the sequencing of viruses, inspired by the need to understand the genomic aspects of the virus responsible for COVID-19. For achieving semantic clarity on the VCM, we leverage the "ontological unpacking" method, a process of ontological analysis that reveals the ontological foundation of the information that is represented in a conceptual model. This is accomplished by applying the stereotypes of the OntoUML ontology-driven conceptual modeling language. As a result, we propose a new OntoVCM, an ontologically grounded model, based on the initial VCM, but with guaranteed interoperability among the data sources that employ it.

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/is

Assessing the value of ontologically unpacking a conceptual model for human genomics

Alberto García S. [a,*,1], Anna Bernasconi [a,b,1], Giancarlo Guizzardi [c], Oscar Pastor [a], Veda C. Storey [d], Ignacio Panach [a]

[a] PROS Research Center & VRAIN Research Institute, Polytechnic University of Valencia, Spain
[b] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Italy
[c] Semantics, Cybersecurity & Services Group (SCS), Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, The Netherlands
[d] J. Mack Robinson College of Business, Georgia State University, United States

ARTICLE INFO

Article history:
Received 11 April 2023
Received in revised form 31 May 2023
Accepted 1 June 2023
Available online 23 June 2023
Recommended by Dennis Shasha

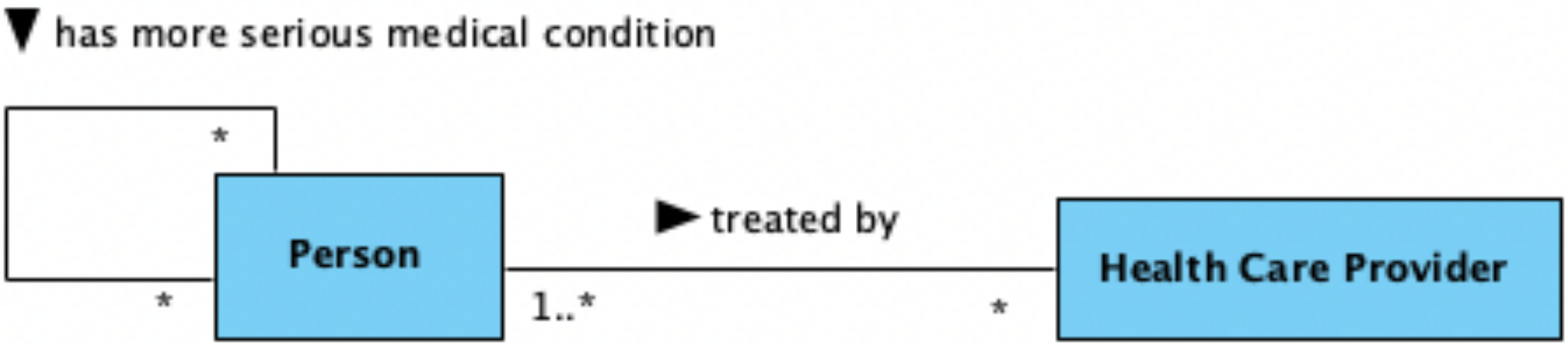
Keywords:
Ontological unpacking
Conceptual modeling
Foundational ontology
OntoUML
Genomics
Metabolic pathways
Data explanation

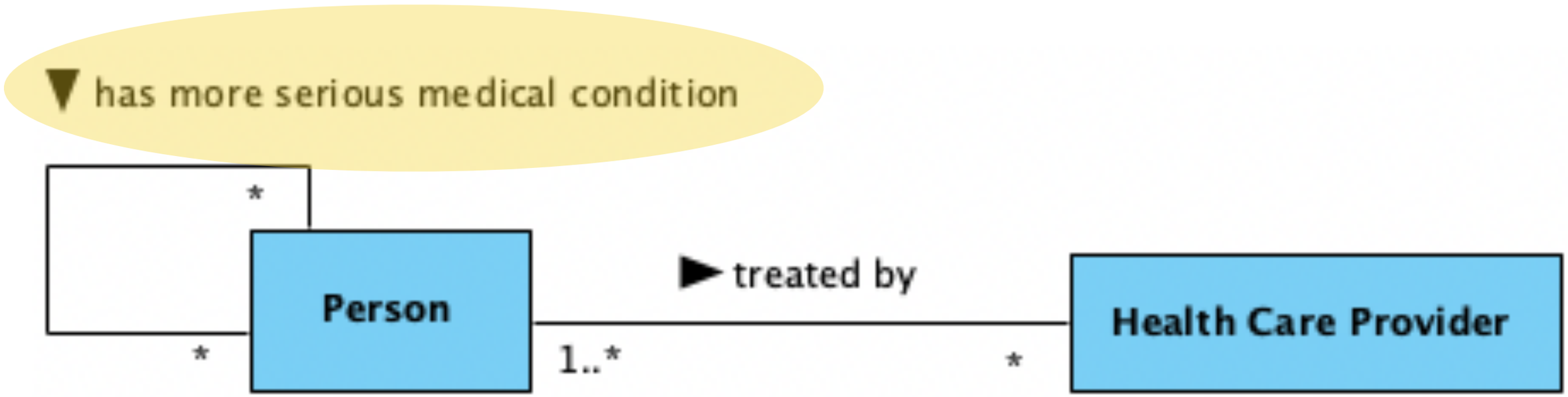
ABSTRACT

Although the knowledge about human genomics is available to all scientists, information about this scientific breakthrough can often be difficult to fully comprehend and share. A Conceptual Schema of the Human Genome was previously developed to assist in describing human genome-related knowledge, by representing a holistic view of the relevant concepts regarding its biology and underlying mechanisms. This model should become helpful for any researcher who works with human genomics data. We, therefore, perform the process of *ontological unpacking* on a portion of the model, to facilitate domain understanding and data exchange among heterogeneous systems. The ontological unpacking is a transformation of an input conceptual model into an enriched model based on a foundational ontology. The preliminary analysis and enrichment process are supported by the ontological conceptual modeling language OntoUML, which has been applied previously to complex models to gain ontological clarity. The value of the used method is first assessed from a theoretical point of view: the transformation results in significant, diverse modeling implications regarding the characterization of biological entities, the representation of their changes over time, and, more specifically, the description of chemical compounds. Since the ontological unpacking process is costly, an empirical evaluation is conducted to study the practical implications of applying it in a real learning setting. A particularly complex domain such as metabolic pathways is either described by adopting a traditional conceptual model or explained through an ontologically unpacked model obtained from a traditional model. Our research is evidence that including a strong ontological foundation in traditional conceptual models is useful. It contributes to designing models that convey biological domains better than the original models.

2

Ontological Unpacking as **Explanation**


has more serious medical condition
*
*
Person
treated by
1..*
*
Health Care Provider


has more serious medical condition
*
Person
*
treated by
1..*
*
Health Care Provider

John

Bob

has-more-serious-medical-condition(Bob,John**)**
iff both Bob and John have medical conditions
and Bob has at least one medical condition
that is more severe than
all the medical conditions of John’s

has-more-serious-medical-condition(Bob,John**)**

John

S1

S2

S3

Bob

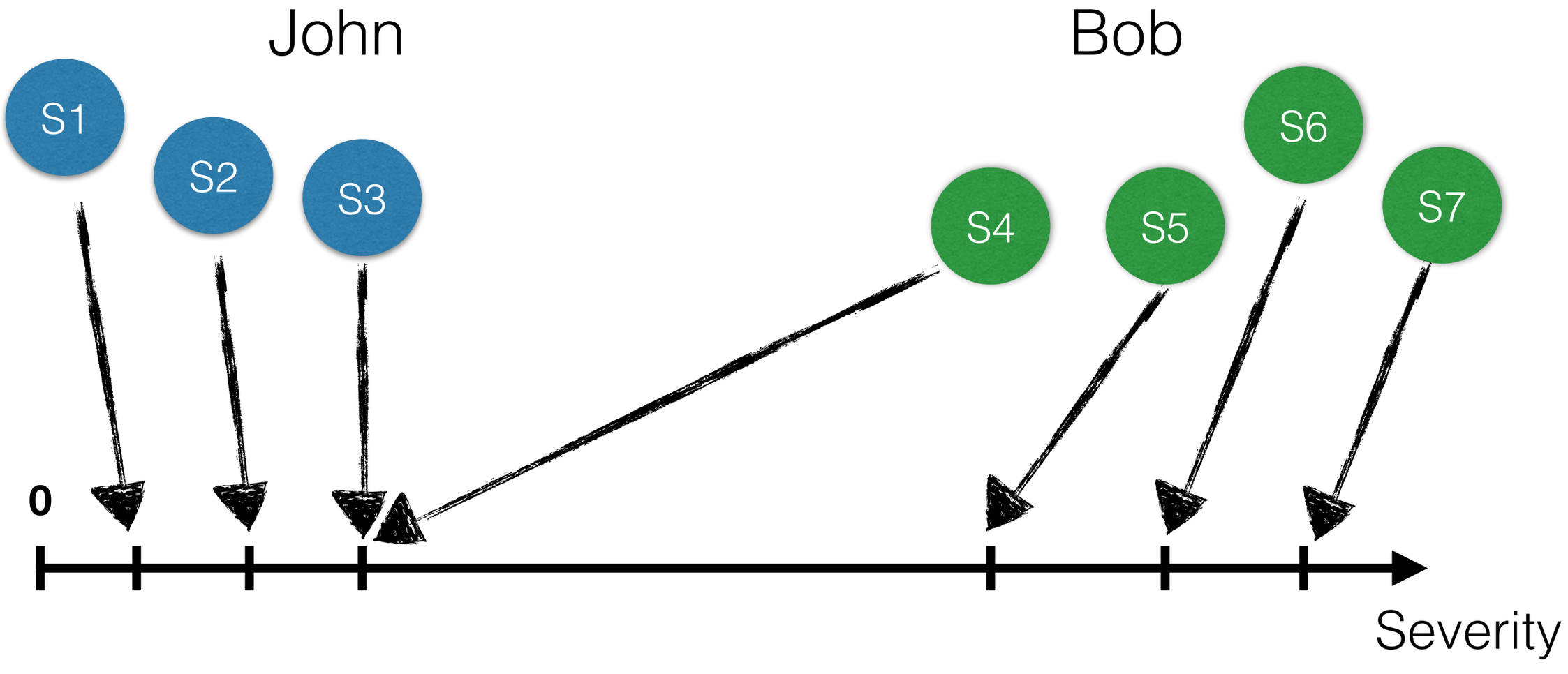
S4

S5

S6

S7

has-more-serious-medical-condition(Bob,John**)**

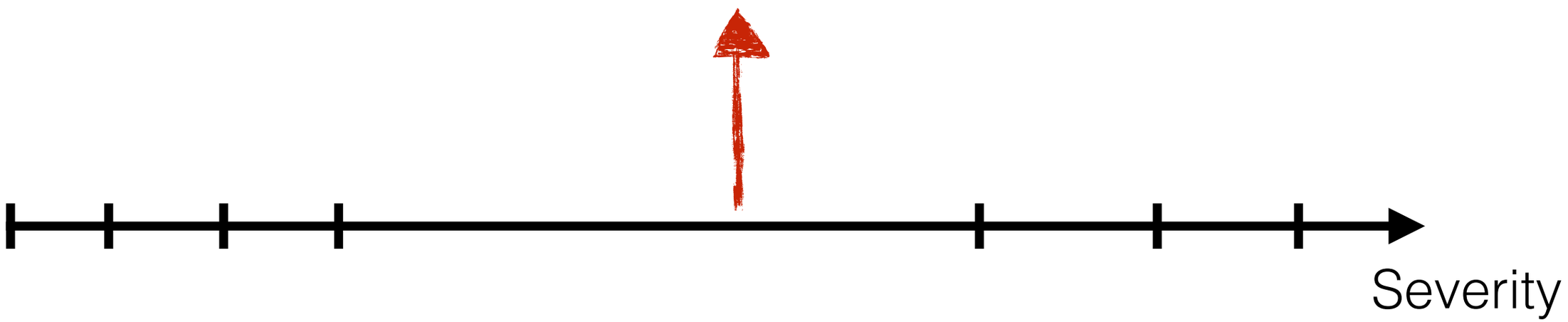


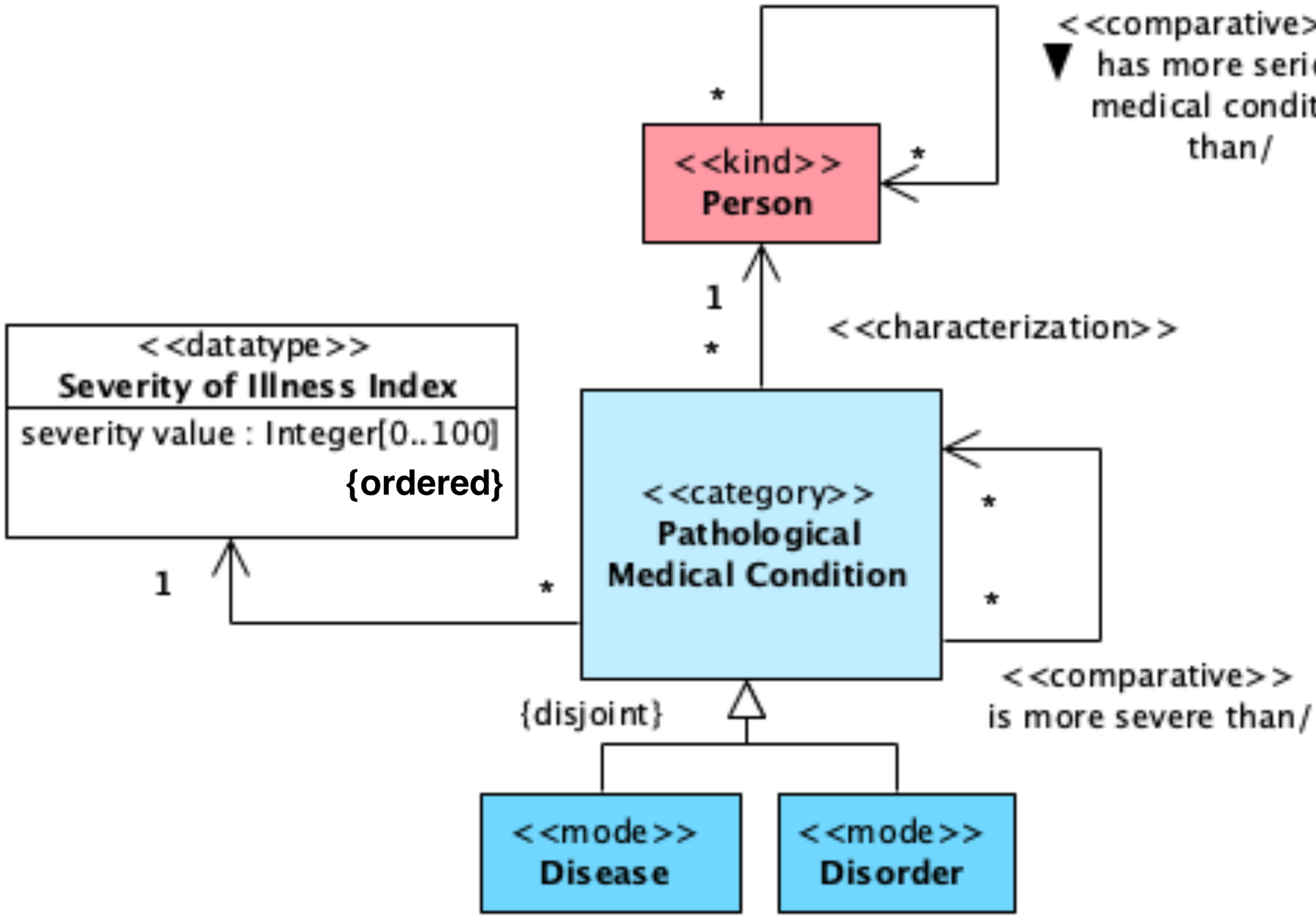
has-more-serious-medical-condition(Bob,John**)**

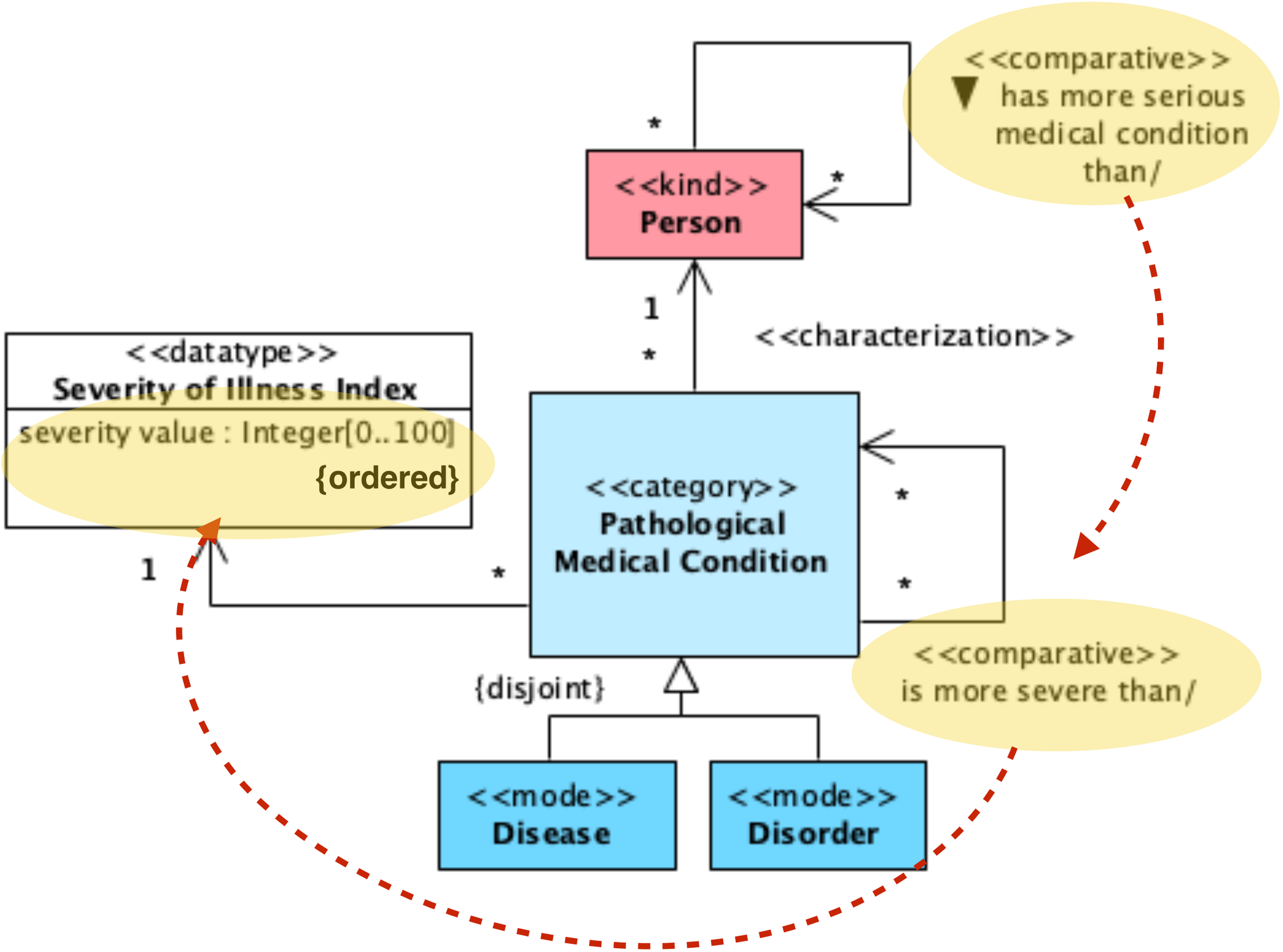
Totally Ordered
(Non-reflexive
Asymmetric
Transitive
Total)

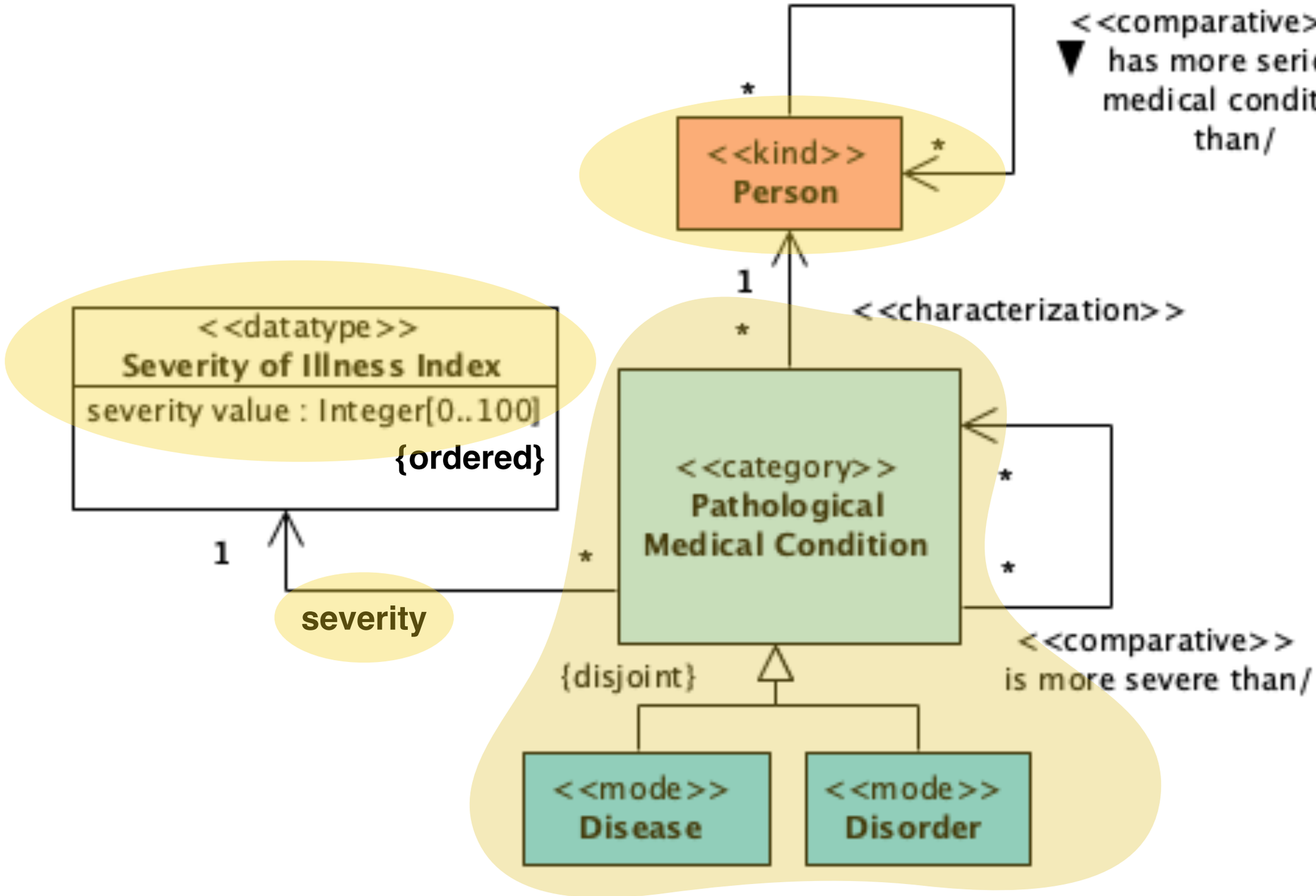
has-more-serious-medical-condition(Bob,John**)**

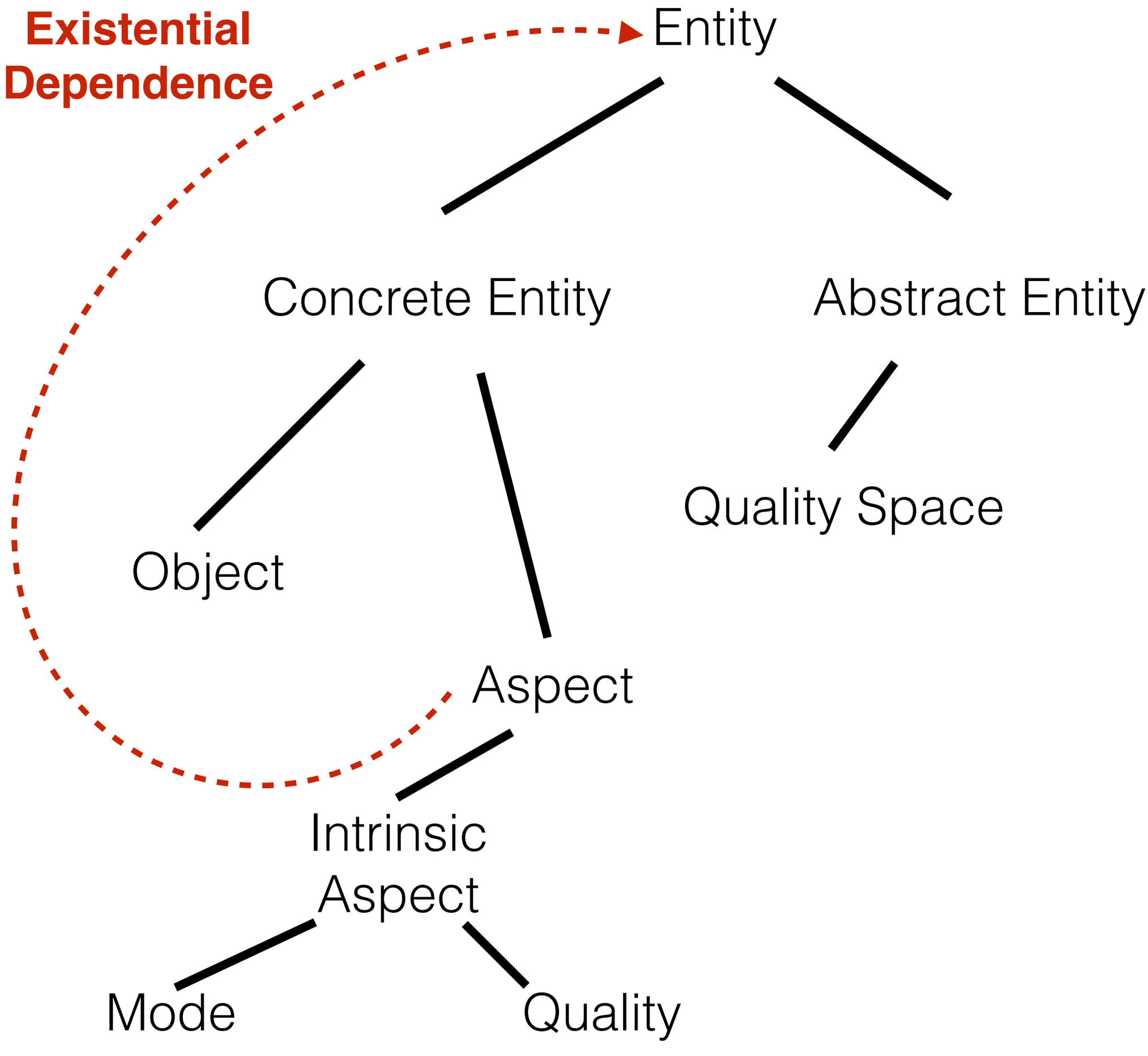
Totally Ordered
(Non-reflexive
Asymmetric
Transitive
Total)

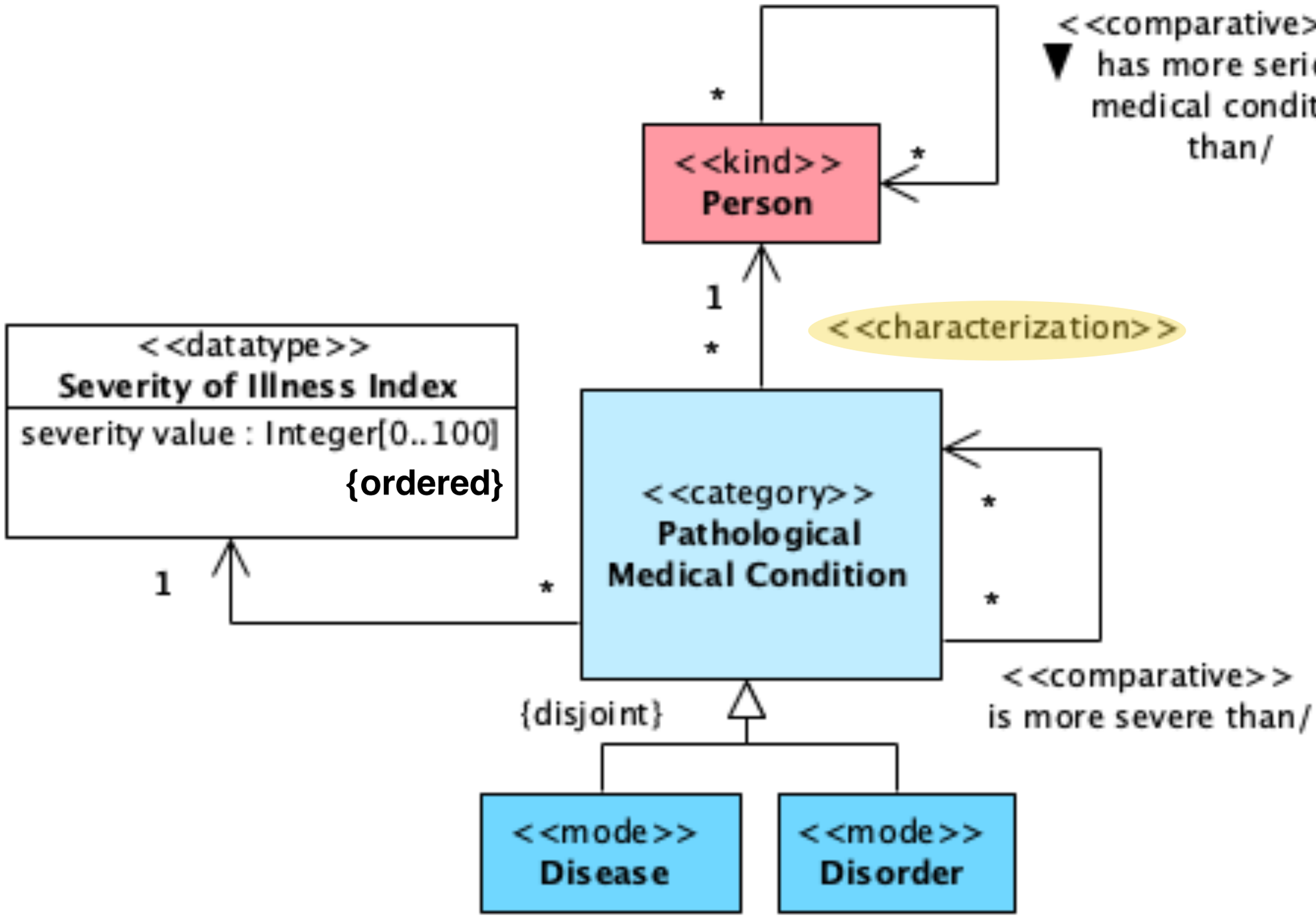


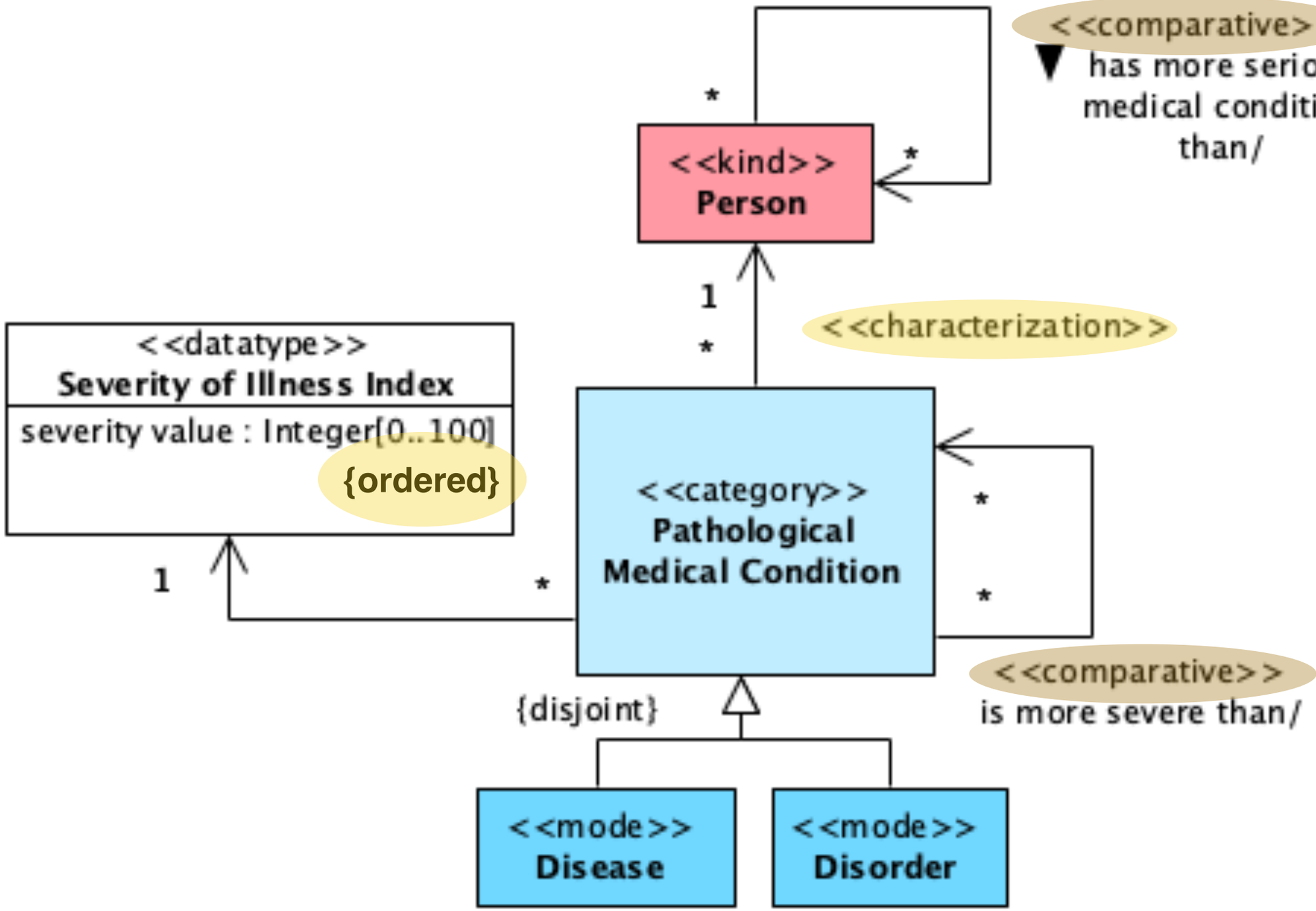

<<comparative>>
has more serious
medical condition
than/
*
*
<<kind>>
Person
1
*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
{ordered}
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder

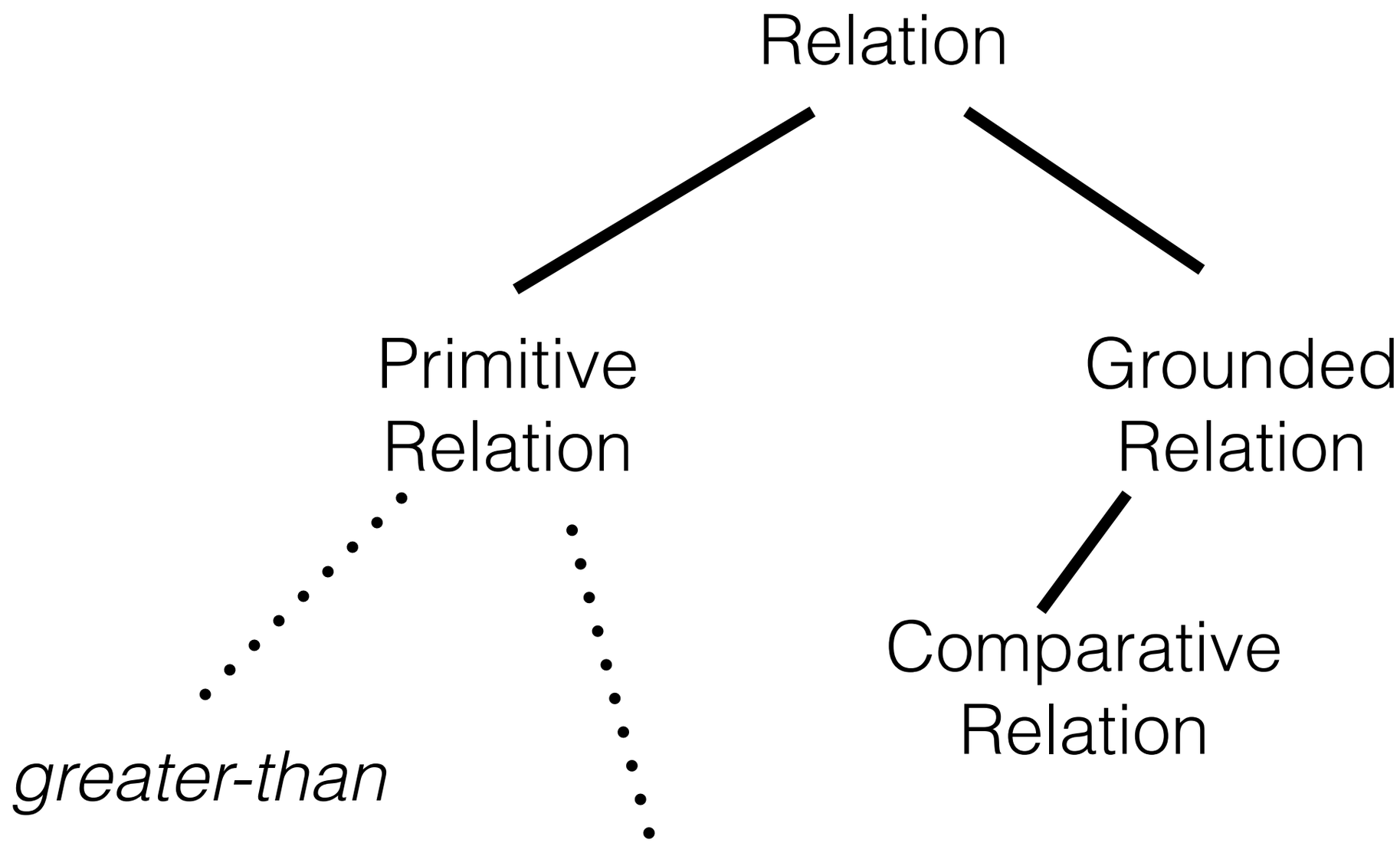

<<comparative>>
has more serious
medical condition
than/
*
*
<<kind>>
Person
1
*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
{ordered}
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder

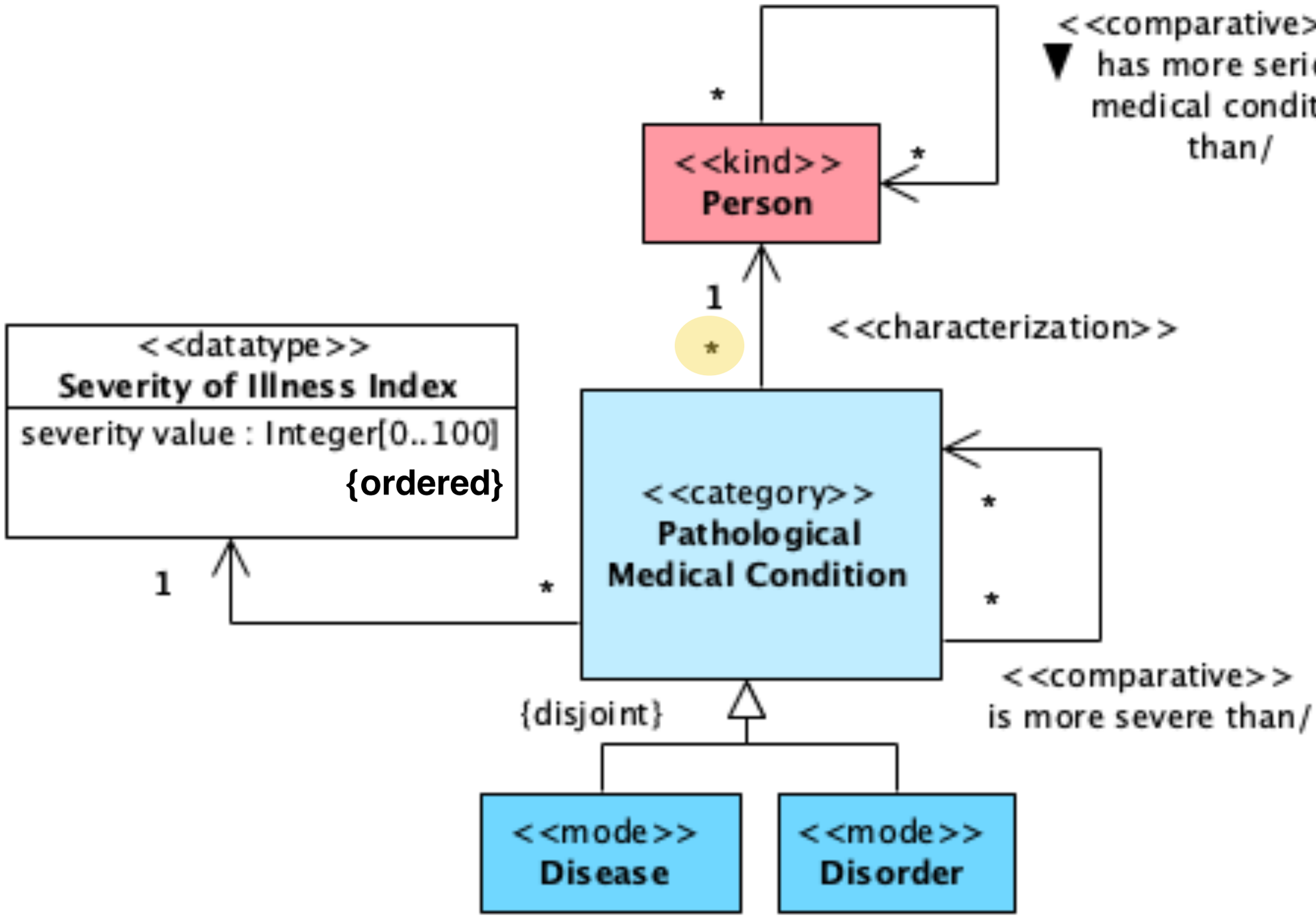

<<comparative>>
has more serious
medical condition
than/
*
*
<<kind>>
Person
1
*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
{ordered}
<<category>>
Pathological
Medical Condition
*
*
1
*
severity
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder

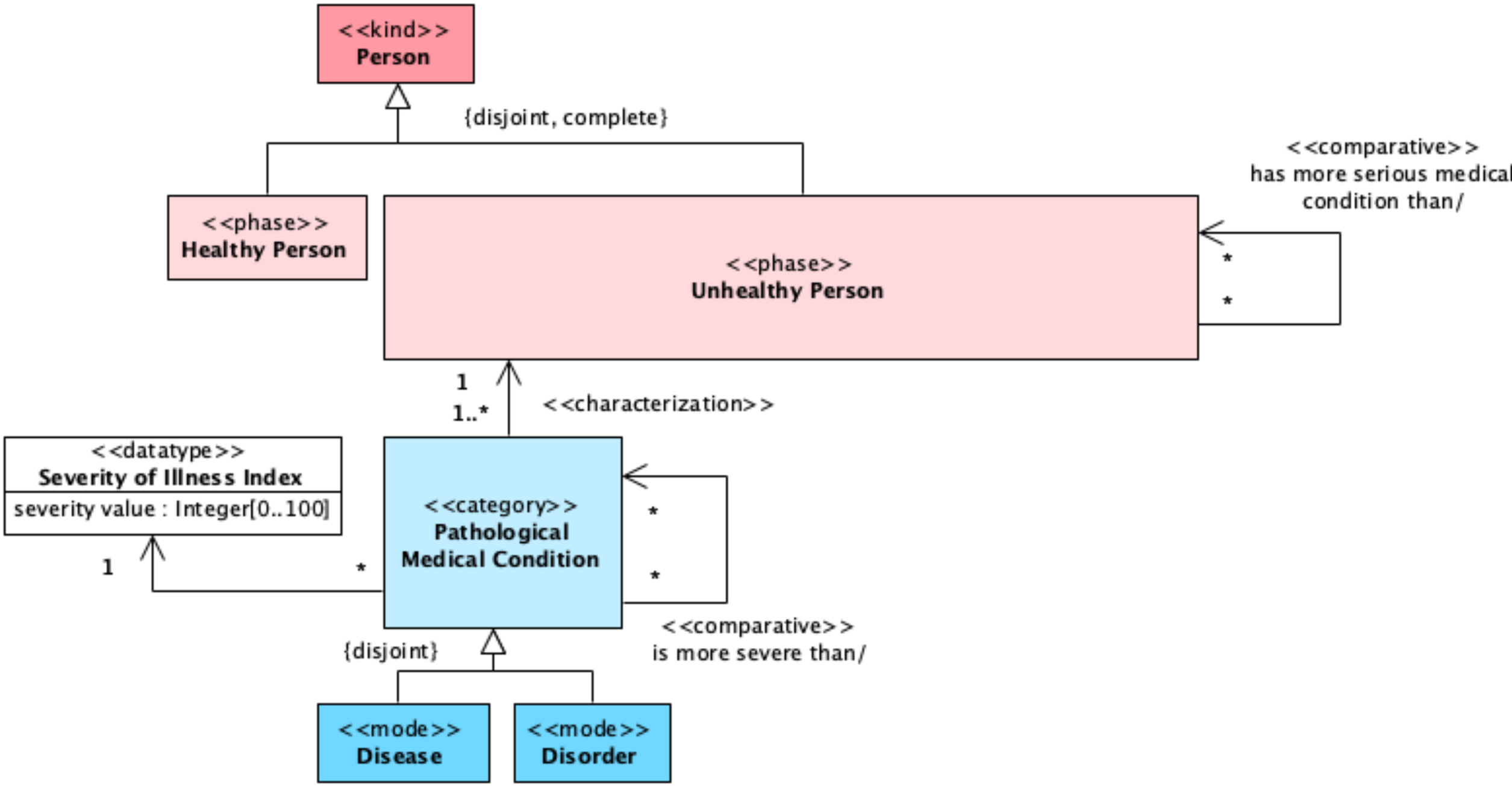

Existential Dependence
Entity
Concrete Entity
Abstract Entity
Quality Space
Object
Aspect
Intrinsic Aspect
Mode
Quality

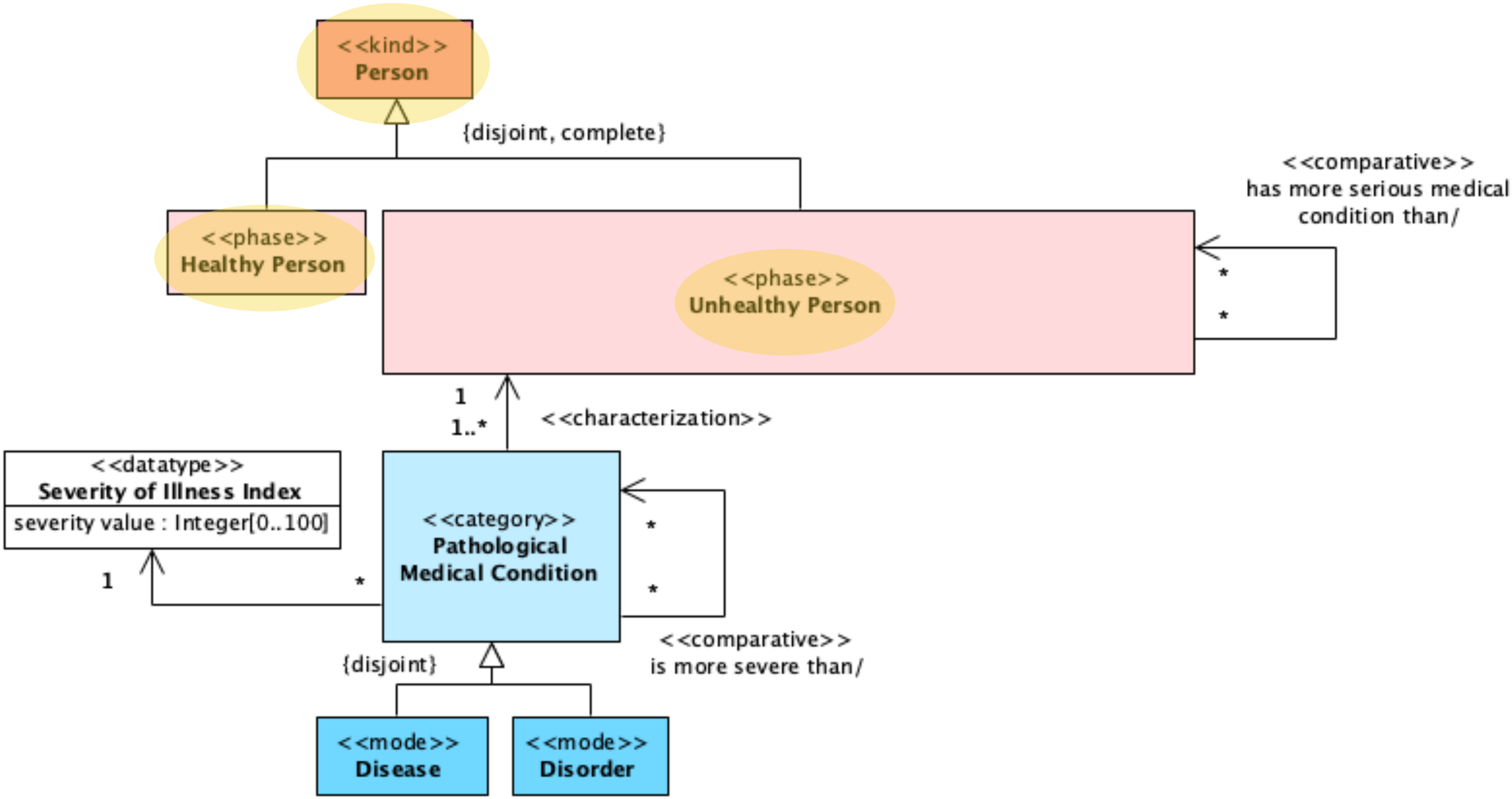

<<comparative>>
has more serious
medical condition
than/
*
*
<<kind>>
Person
1
*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
{ordered}
<<category>>
Pathological
Medical Condition
1
*
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder

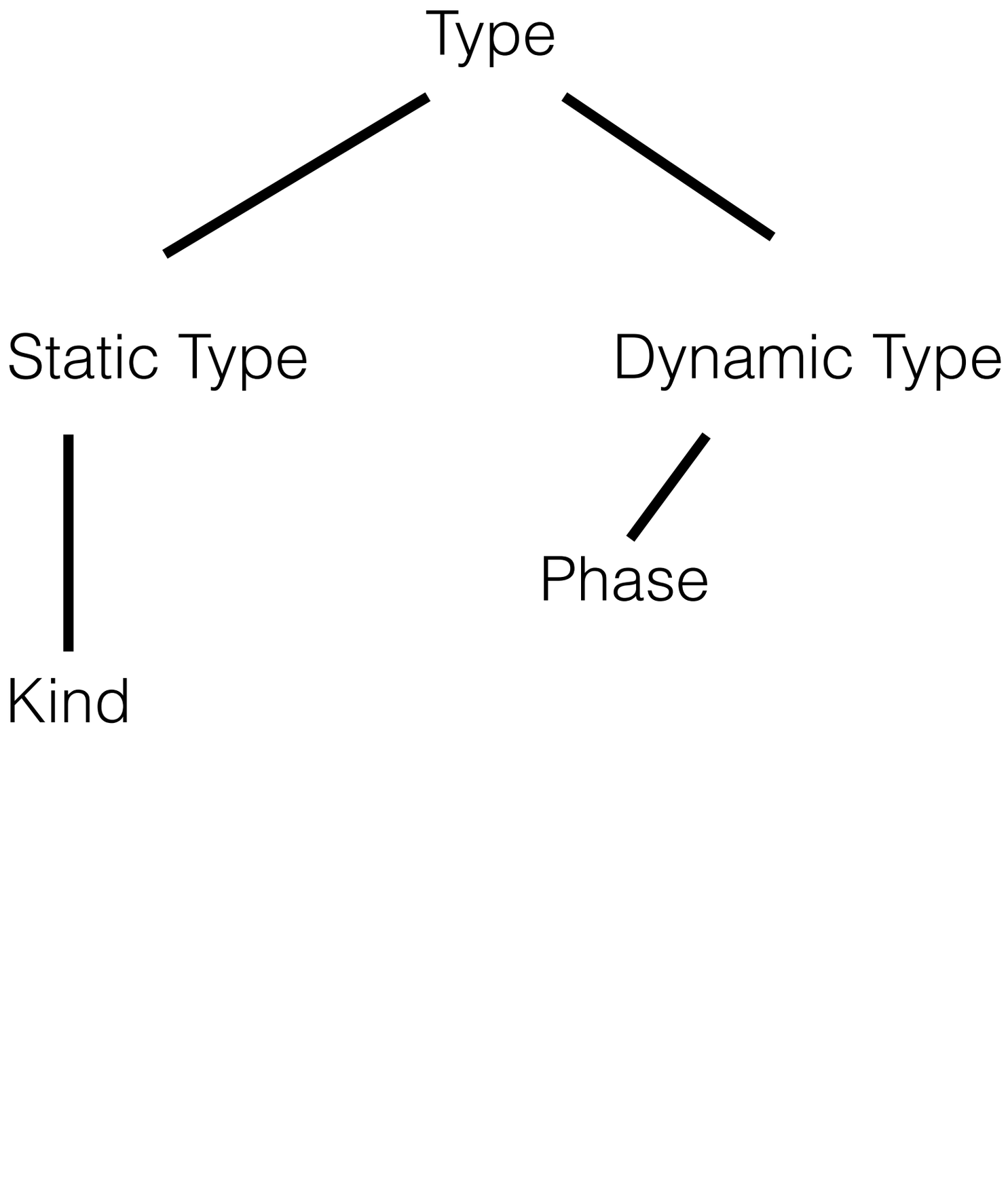

<<comparative>>
has more serious medical condition than/
*
*
<<kind>>
Person
1
*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
{ordered}
<<category>>
Pathological Medical Condition
*
*
1
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder

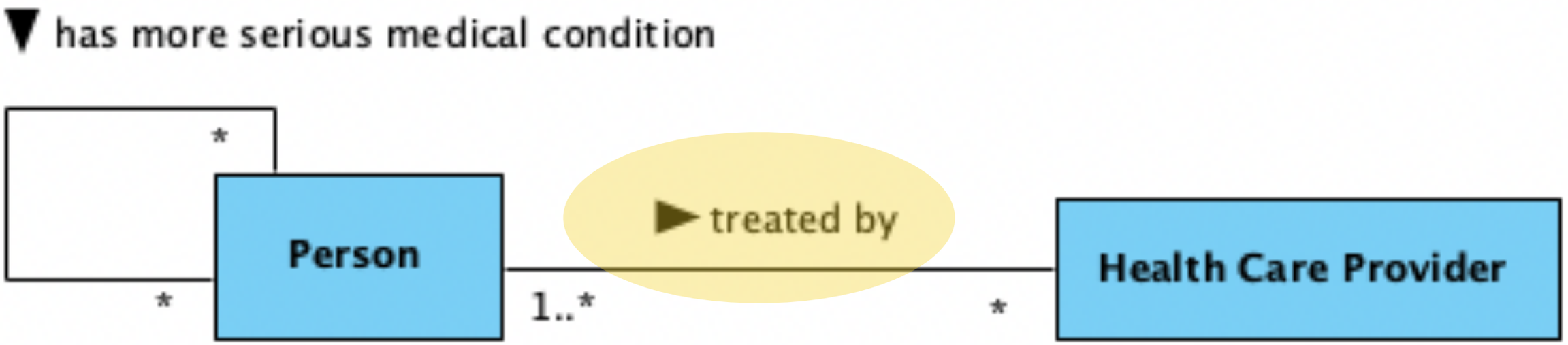

Relation
Primitive Relation
Grounded Relation
Comparative Relation
greater-than
characterization
(Existential Dependence)


<<comparative>>
has more serious
medical condition
than/
*
*
<<kind>>
Person
1
*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
{ordered}
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder


<<kind>>
Person
{disjoint, complete}
<<phase>>
Healthy Person
<<phase>>
Unhealthy Person
<<comparative>>
has more serious medical condition than/
*
*
1
1..*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
1
*
<<category>>
Pathological Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder


<<kind>>
Person
{disjoint, complete}
<<phase>>
Healthy Person
<<phase>>
Unhealthy Person
<<comparative>>
has more serious medical
condition than/
*
*
1
1..*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder


Type
Static Type
Dynamic Type
Kind
Phase


has more serious medical condition
*
*
Person
1..*
treated by
*
Health Care Provider

Relation

Primitive Relation

greater-than

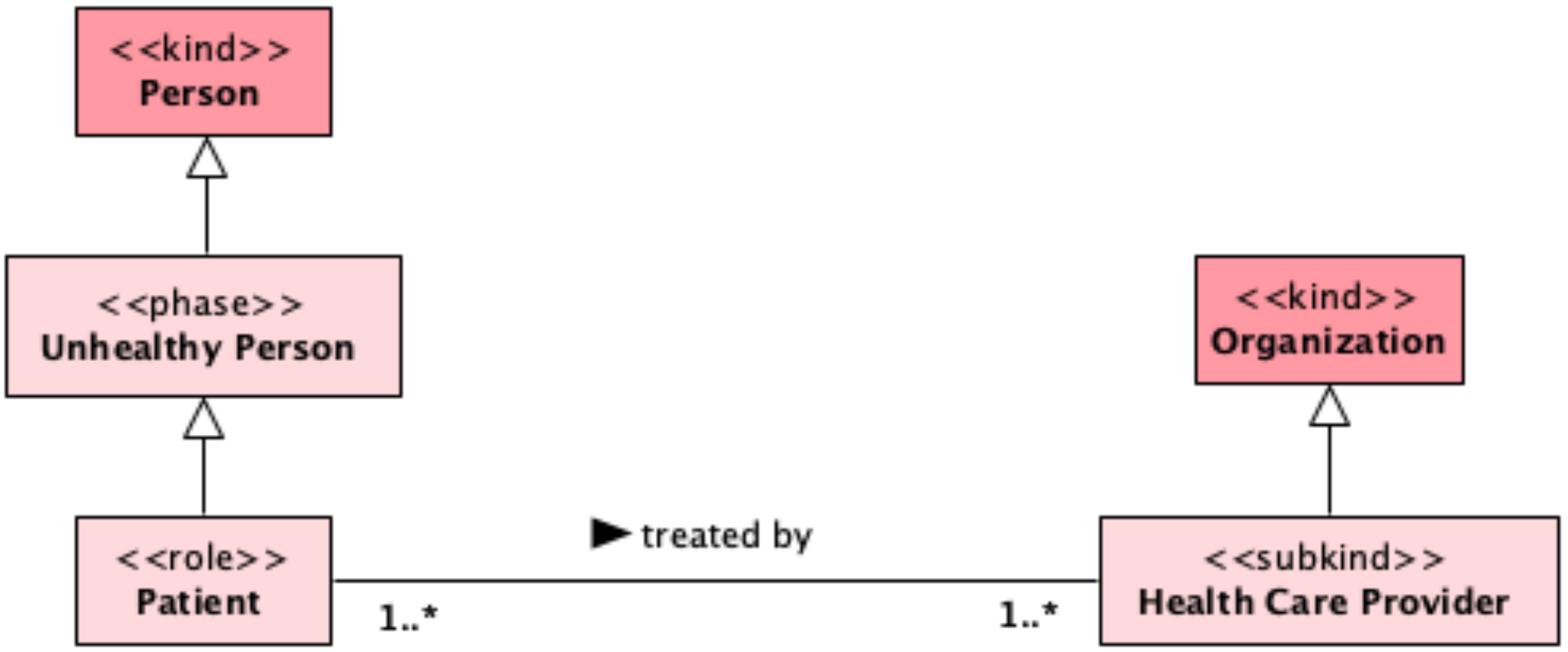
Grounded Relation

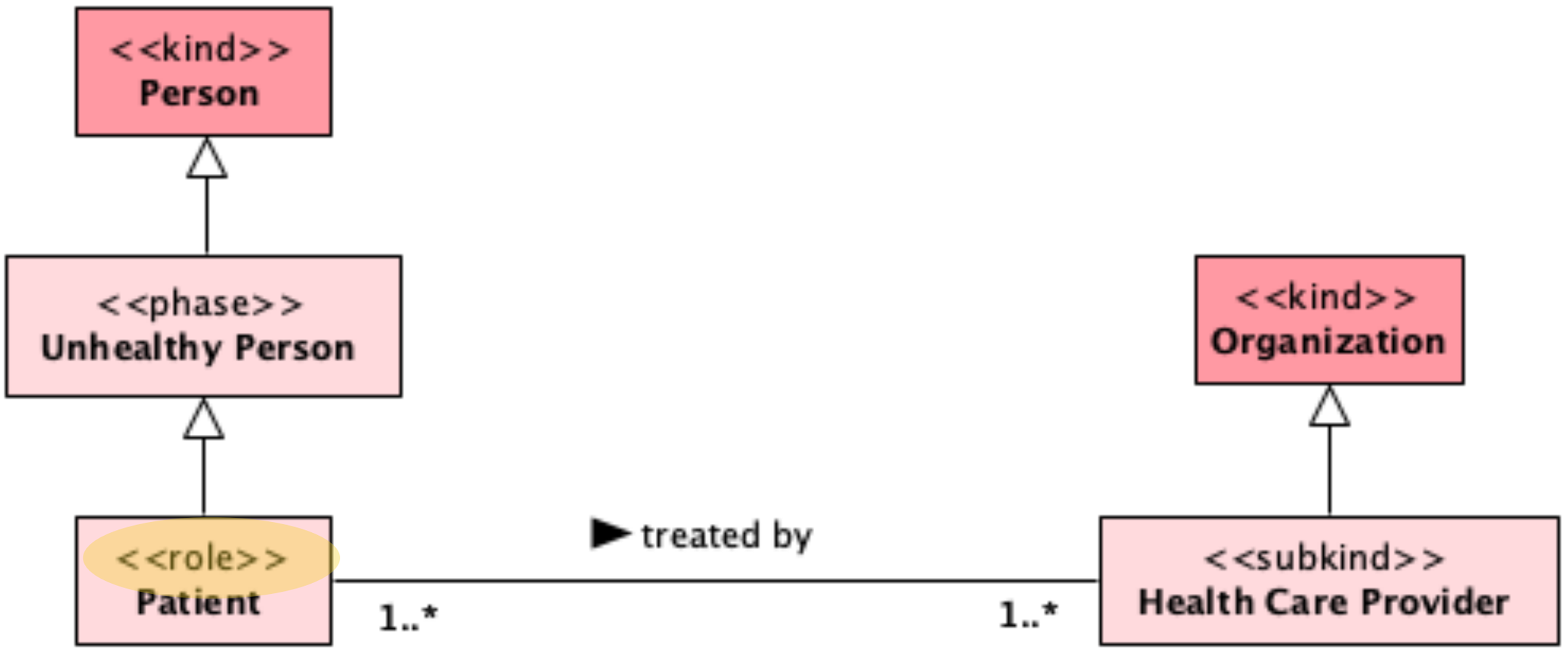
Comparative Relation

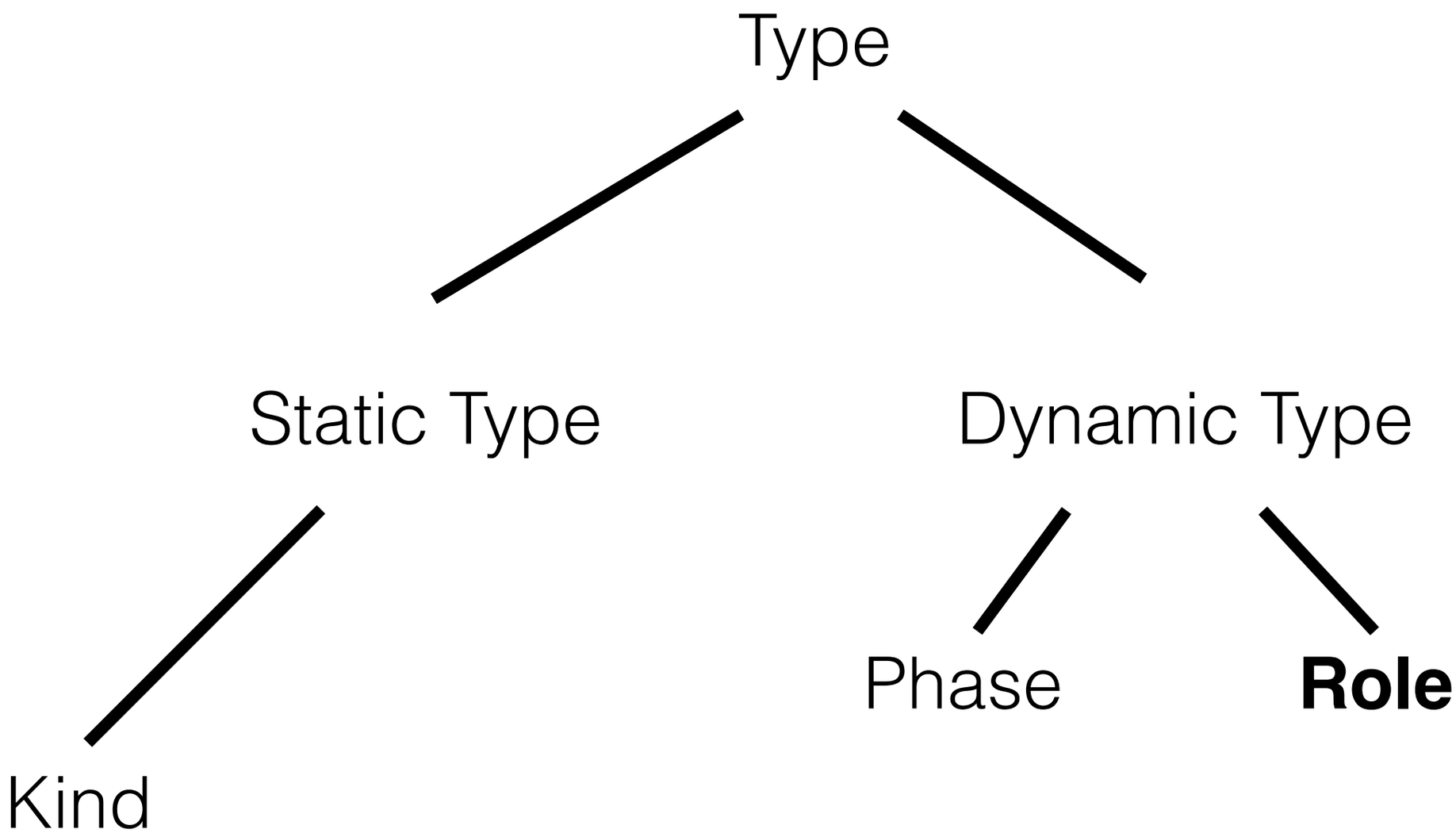
Material Relation

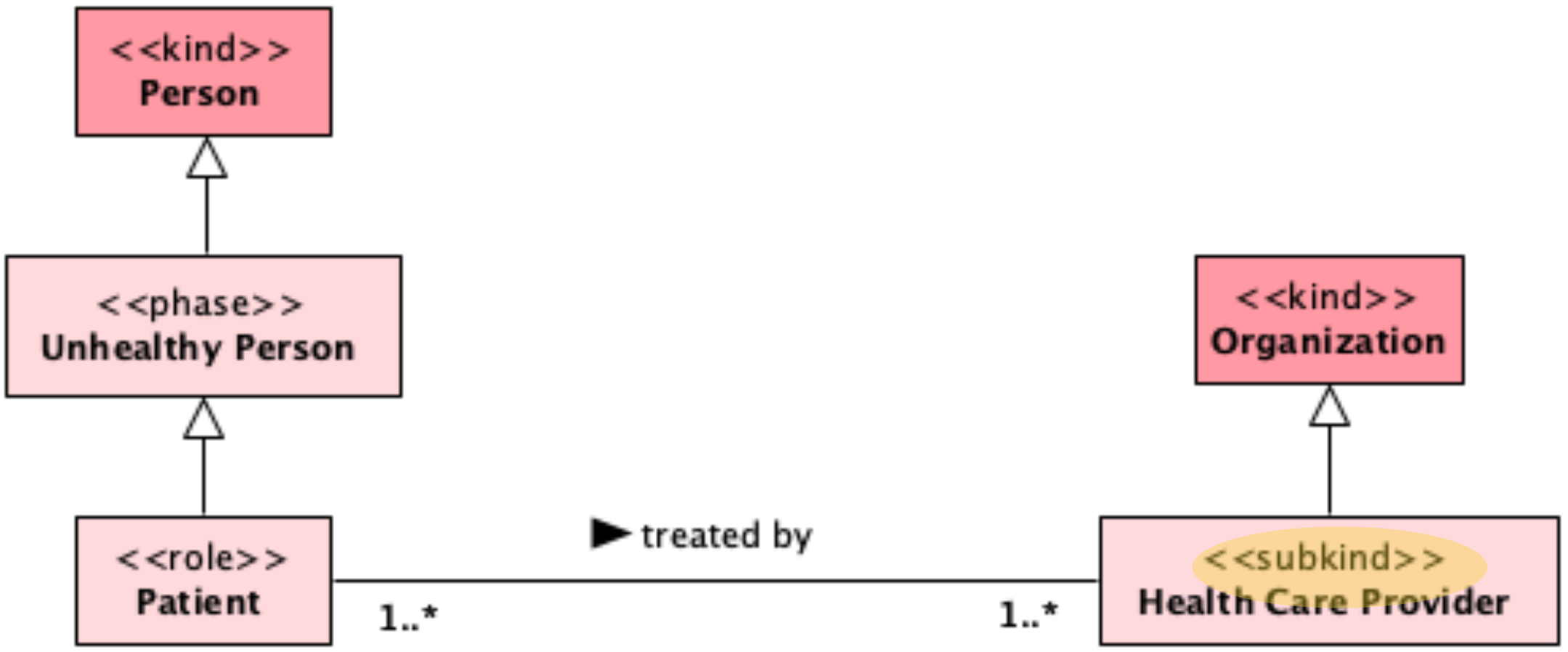
characterization
(Existential Dependence)

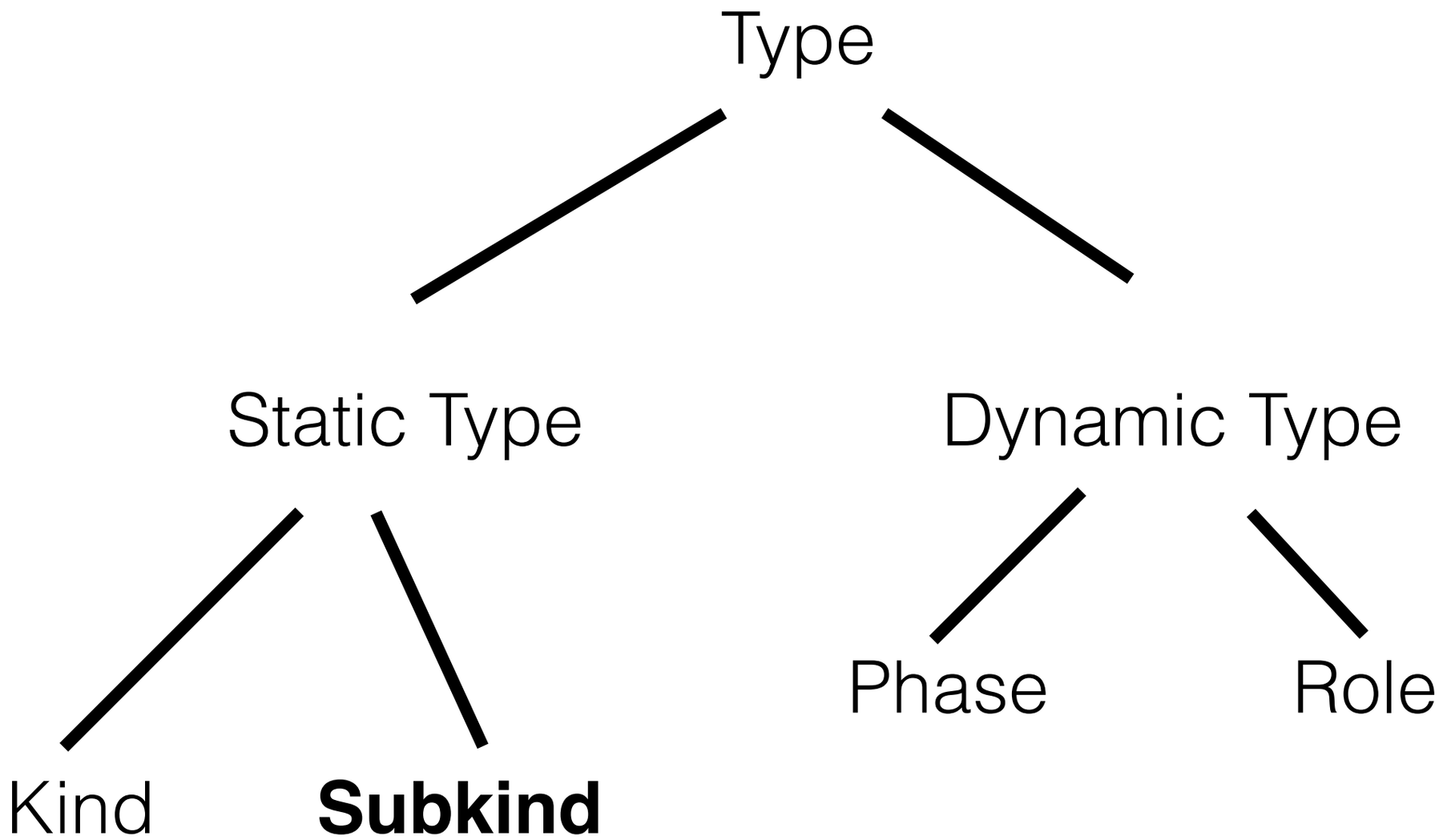

has more serious medical condition
*
*
Person
treated by
1..*
*
Health Care Provider


<<kind>>
Person
<<phase>>
Unhealthy Person
<<role>>
Patient
treated by
1..*
1..*
<<kind>>
Organization
<<subkind>>
Health Care Provider


<<kind>>
Person
<<phase>>
Unhealthy Person
<<role>>
Patient
treated by
1..*
1..*
<<kind>>
Organization
<<subkind>>
Health Care Provider


Type
Static Type
Dynamic Type
Kind
Phase
Role

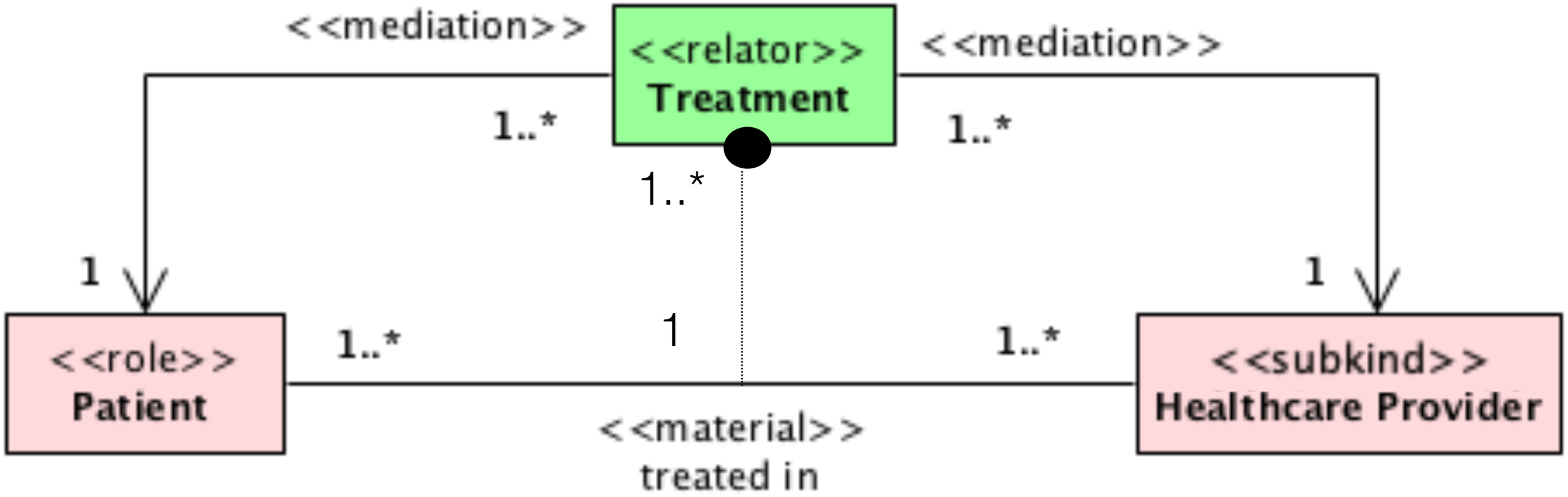

<<kind>>
Person
<<phase>>
Unhealthy Person
<<role>>
Patient
treated by
1..*
1..*
<<kind>>
Organization
<<subkind>>
Health Care Provider

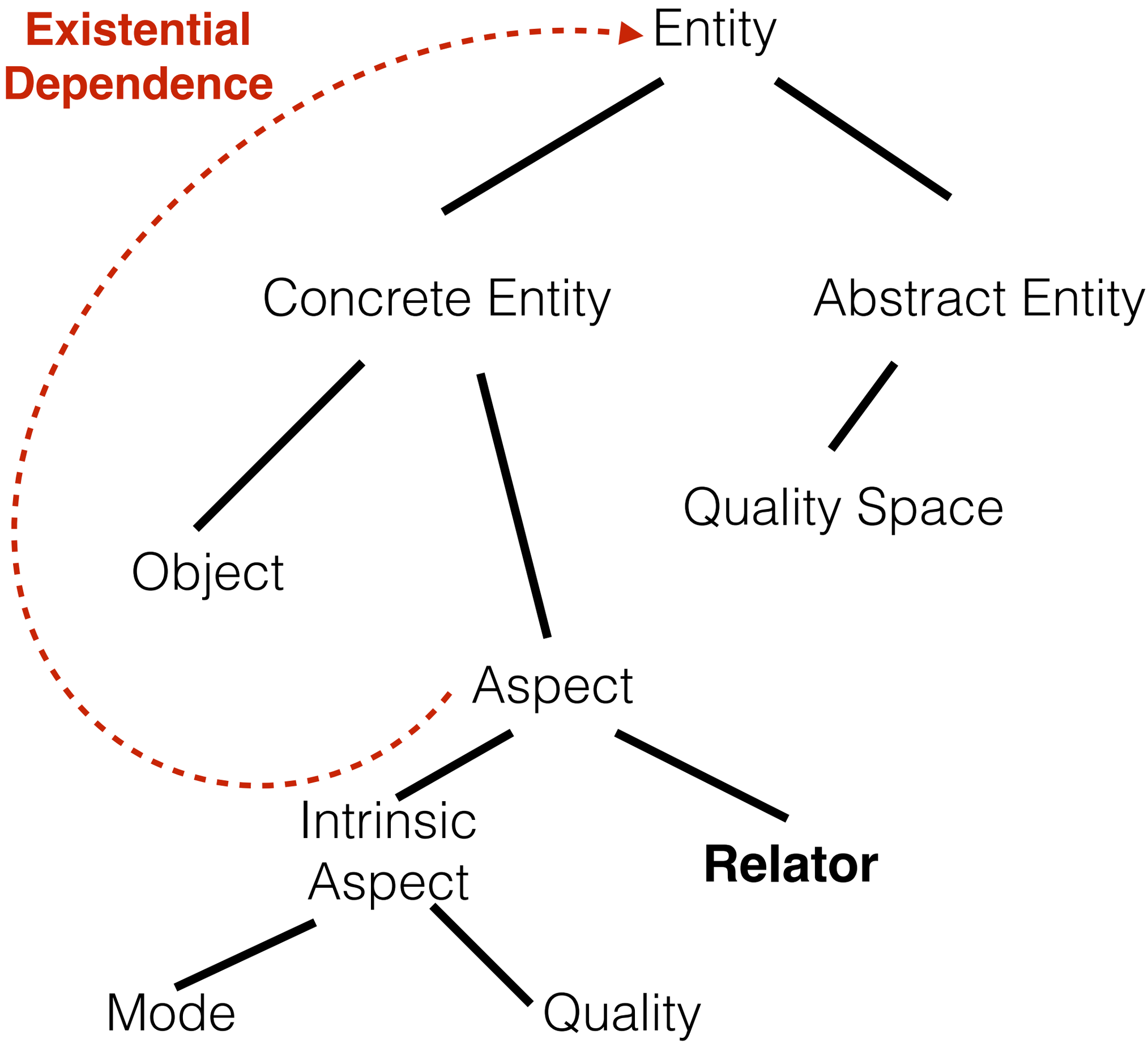

Type
Static Type
Dynamic Type
Kind
Subkind
Phase
Role

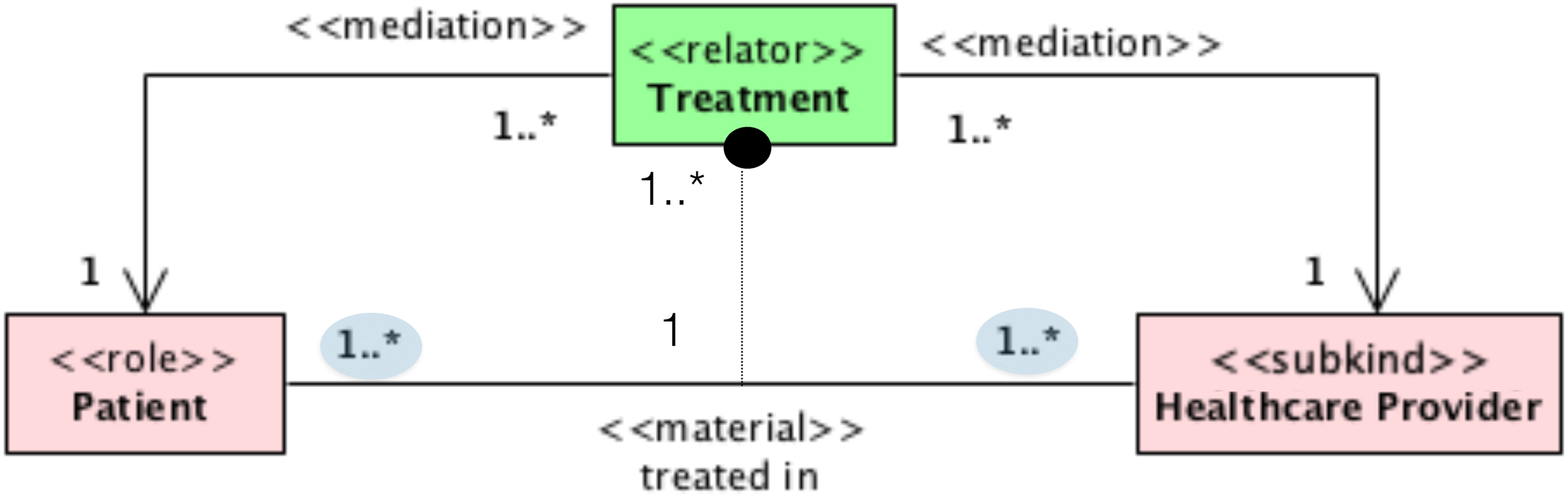

<<kind>>
Person
<<phase>>
Unhealthy Person
<<role>>
Patient
treated by
1..*
1..*
<<kind>>
Organization
<<subkind>>
Health Care Provider

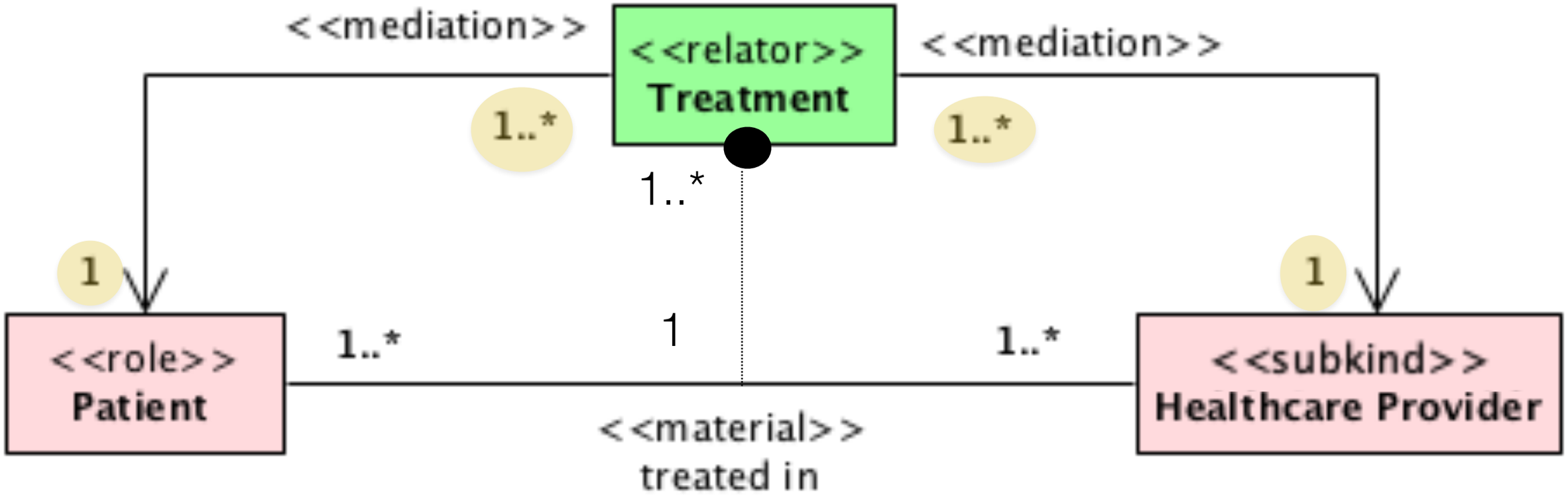
How to interpret 1..* ?

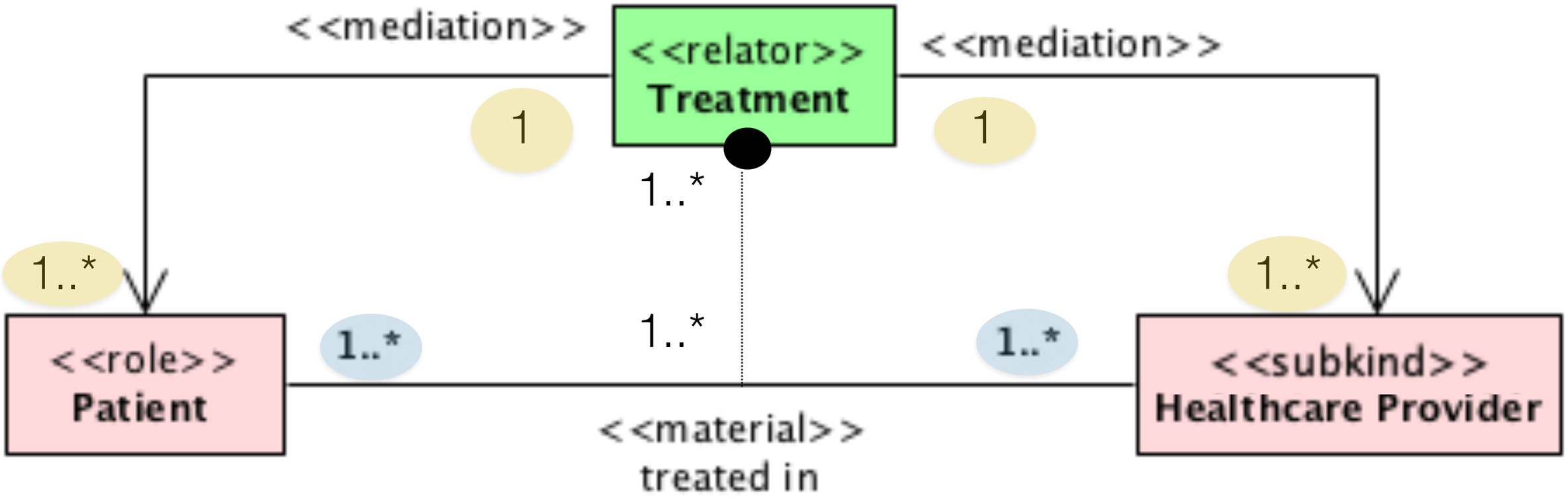
- Given a treatment, there is **exactly one patient**, **exactly one healthcare provider** but **both** patient and healthcare provider **can participate in many treatments**
- Given a treatment, there are **possibly many patients**, **exactly one healthcare provider** but **both** patient and healthcare provider **can participate in many treatments**
- Given a treatment, there is **exactly one patient, possibly many healthcare providers** but **both** patient and healthcare provider **can participate in many treatments**
- Given a treatment, there are **possibly many patients**, **possibly many healthcare providers** and **both** patient and healthcare provider **can participate in many treatments**
- Given a treatment, there are **possibly many patients**, **possibly many healthcare providers** and **both** patient and healthcare provider **can participate in exactly one treatment**
- …

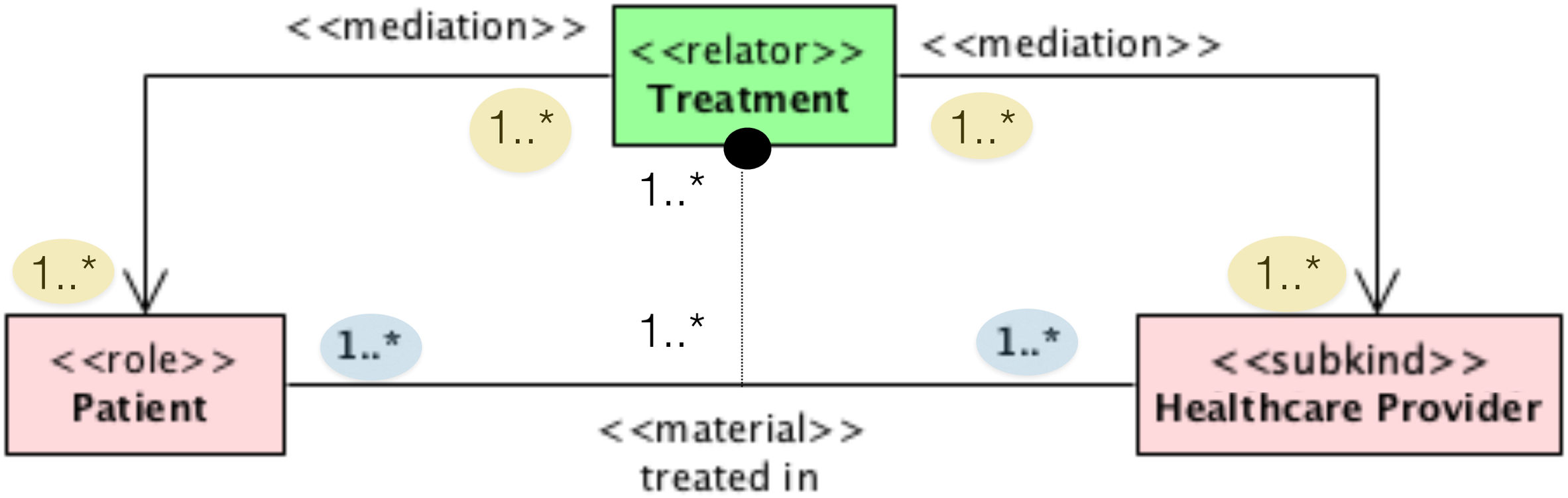

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1
1
1..*
<<role>>
Patient
<<material>>
treated in
<<subkind>>
Healthcare Provider

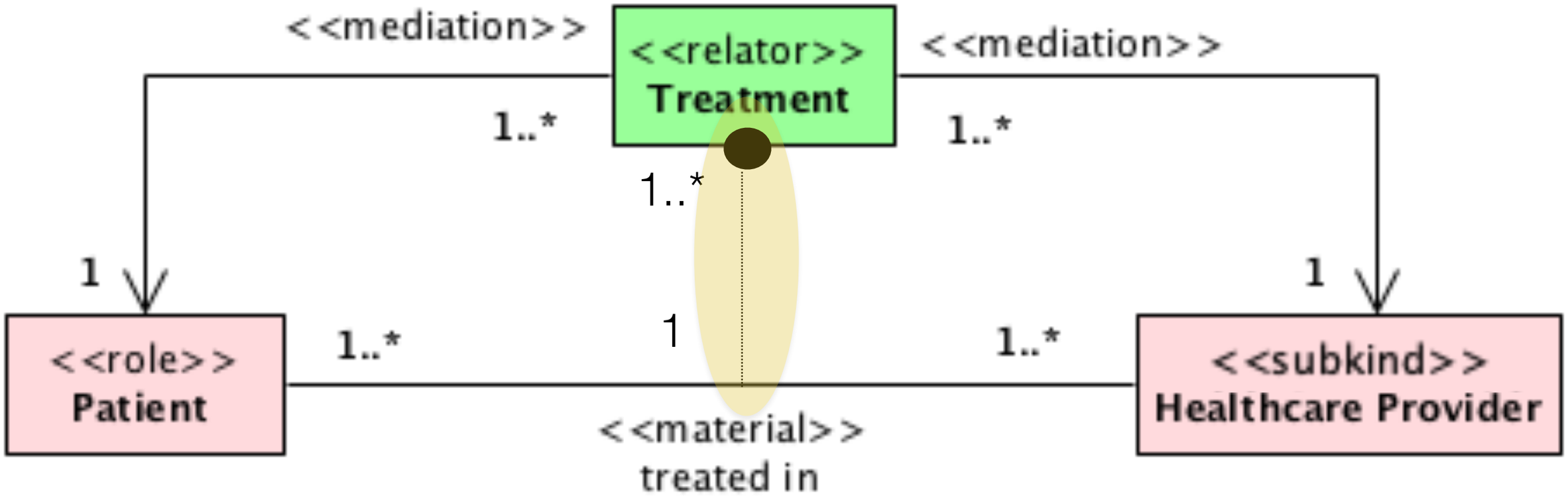

Existential Dependence
Entity
Concrete Entity
Abstract Entity
Quality Space
Object
Aspect
Intrinsic Aspect
Relator
Mode
Quality

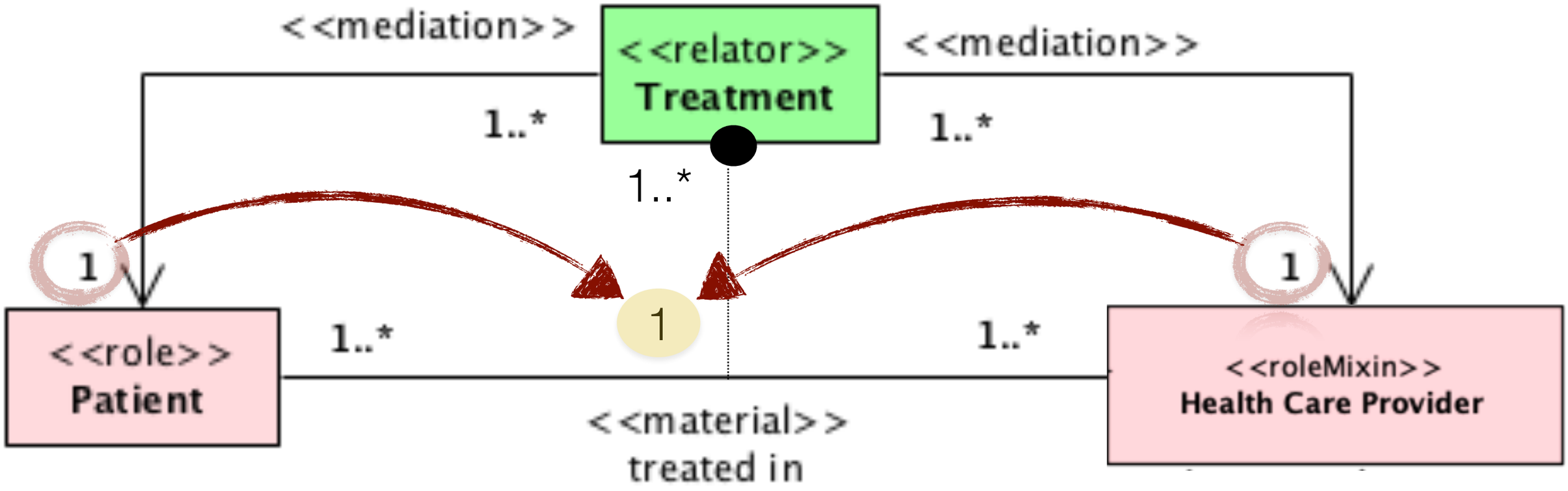

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1
1
1..*
1..*
<<role>>
Patient
<<subkind>>
Healthcare Provider
<<material>>
treated in

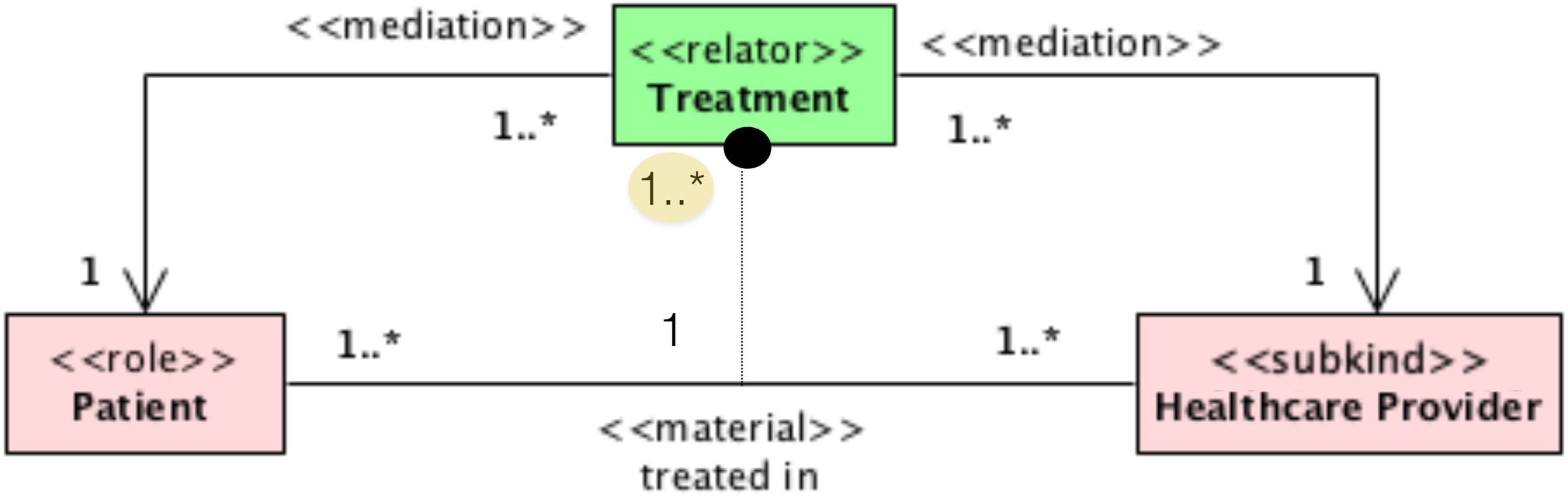

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1
1
1..*
<<role>>
Patient
<<subkind>>
Healthcare Provider
<<material>>
treated in

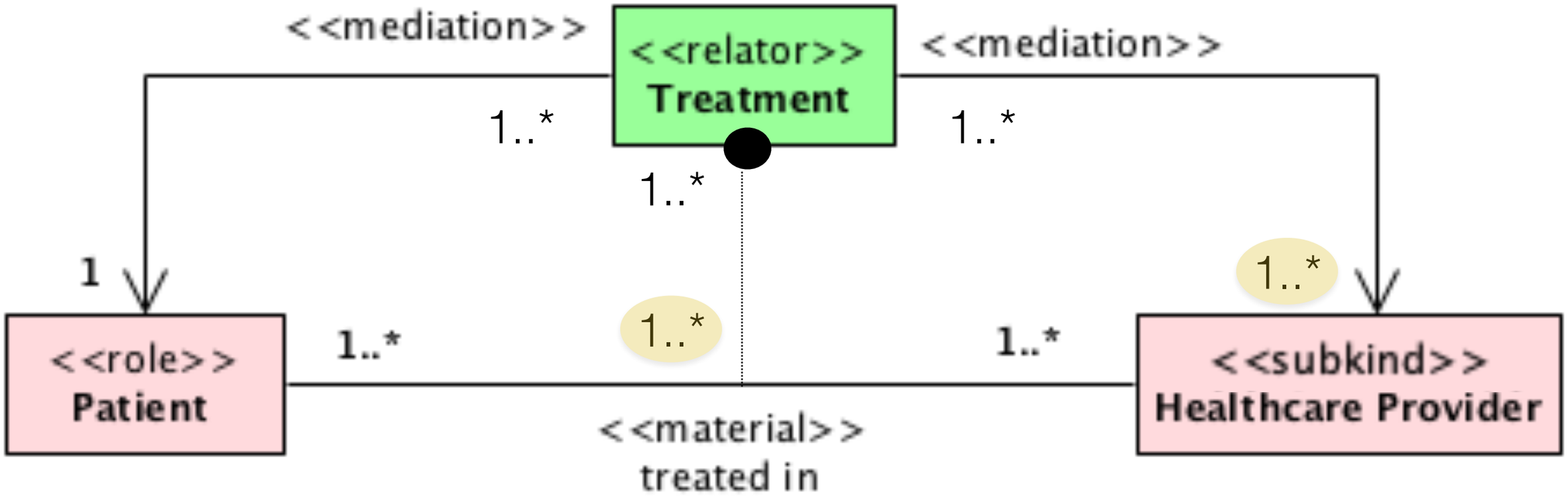

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1
1
1..*
1..*
1..*
1..*
<<role>>
Patient
<<subkind>>
Healthcare Provider
1..*
<<material>>
treated in

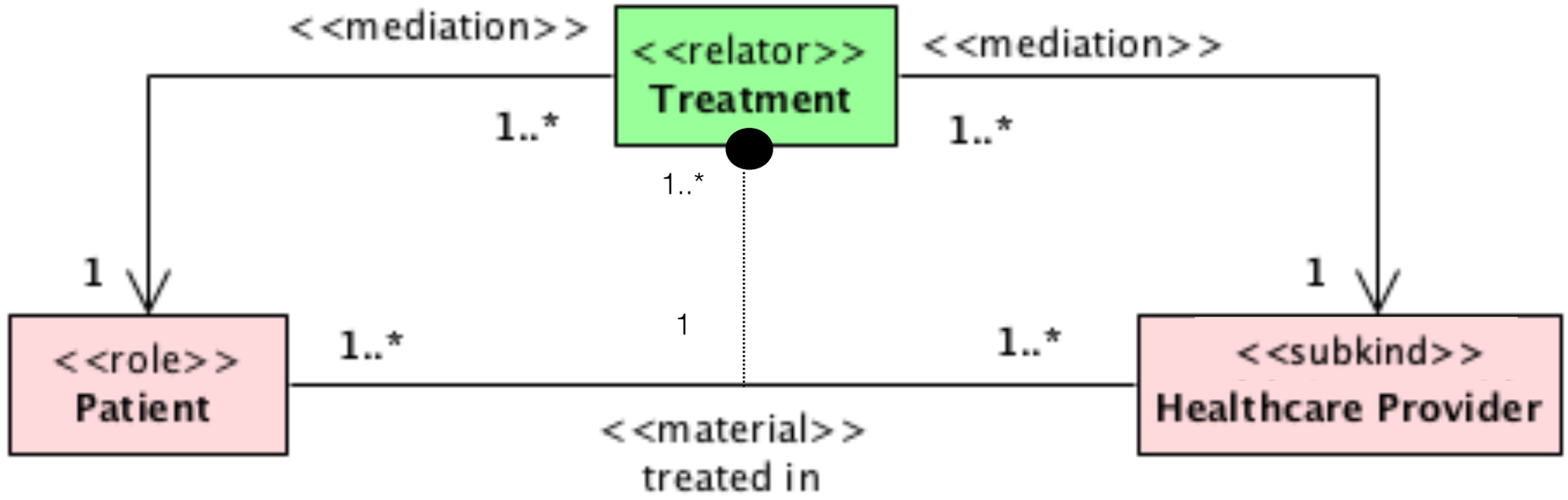

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1..*
<<role>>
Patient
1..*
1..*
1..*
<<subkind>>
Healthcare Provider
<<material>>
treated in

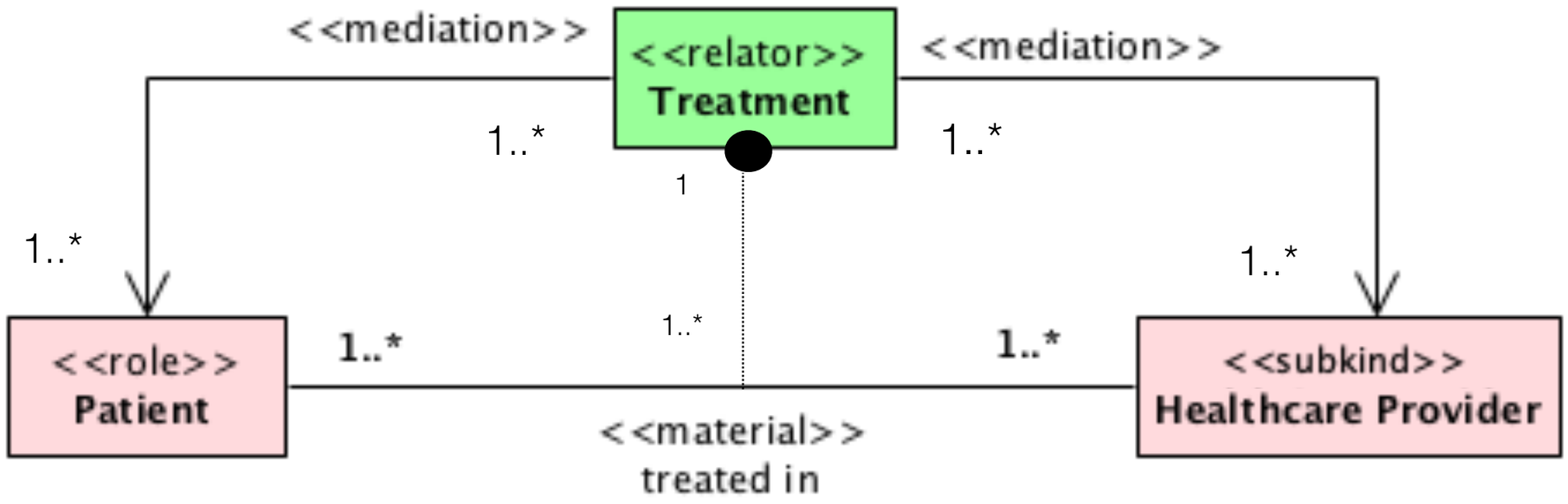

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1
1
1..*
<<role>>
Patient
<<material>>
treated in
<<subkind>>
Healthcare Provider

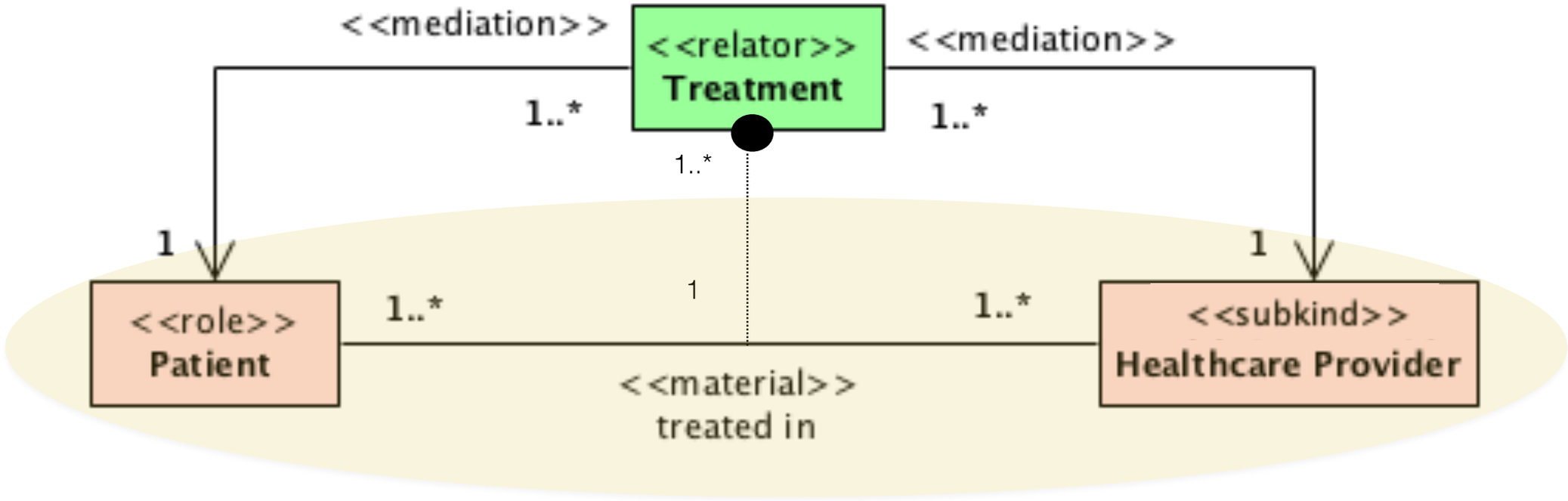

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1
1
1
1..*
1..*
<<role>>
Patient
<<roleMixin>>
Health Care Provider
<<material>>
treated in

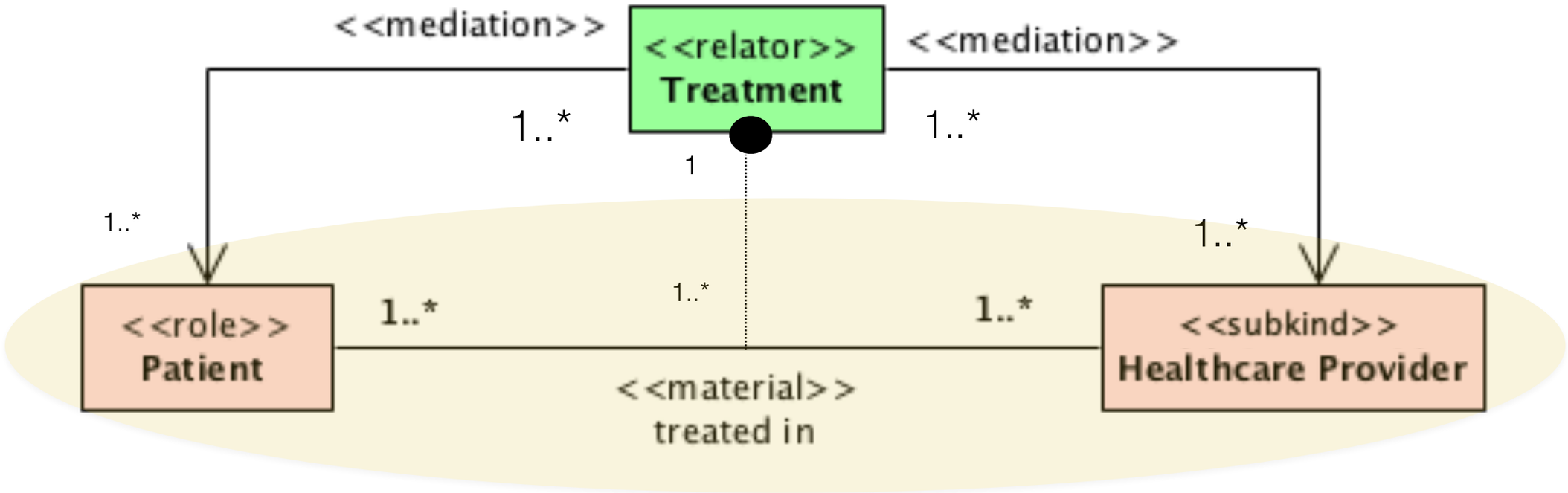

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1
1
1..*
<<role>>
Patient
<<subkind>>
Healthcare Provider
<<material>>
treated in

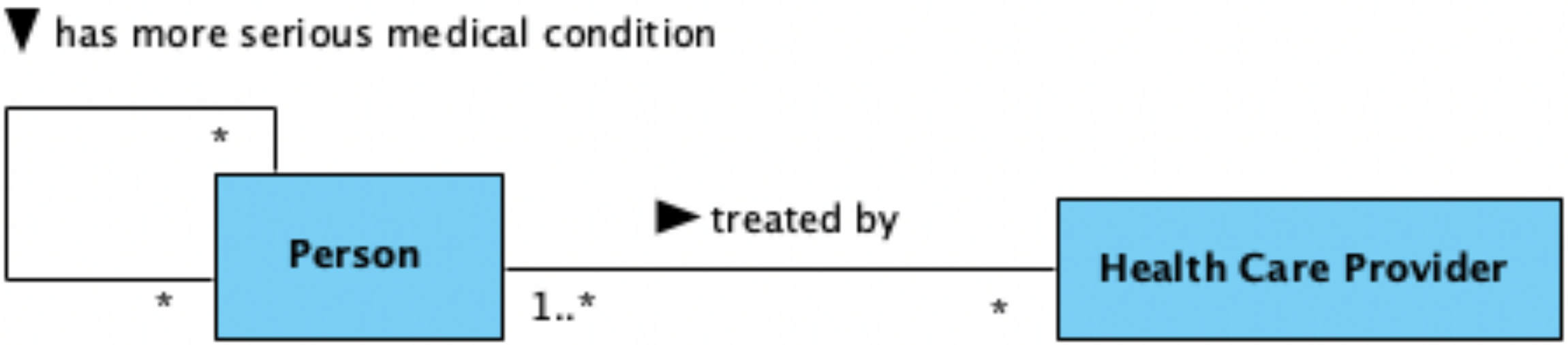

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1
1..*
1..*
1..*
<<role>>
Patient
<<subkind>>
Healthcare Provider
<<material>>
treated in

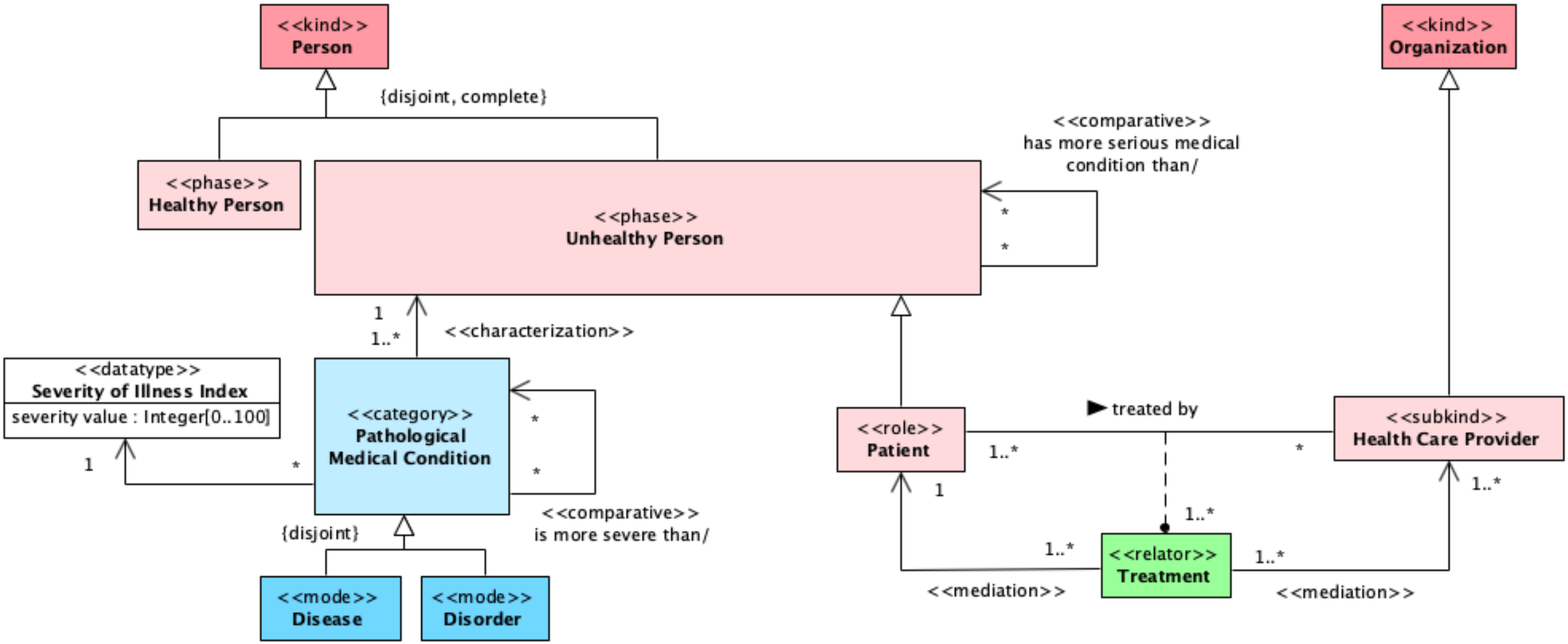

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1
1
<<role>>
Patient
1..*
1..*
<<subkind>>
Healthcare Provider
<<material>>
treated in


<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1
1..*
1..*
<<role>>
Patient
1..*
1..*
<<subkind>>
Healthcare Provider
<<material>>
treated in


<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1..*
1
1
<<role>>
Patient
1..*
1
1..*
<<subkind>>
Healthcare Provider
<<material>>
treated in

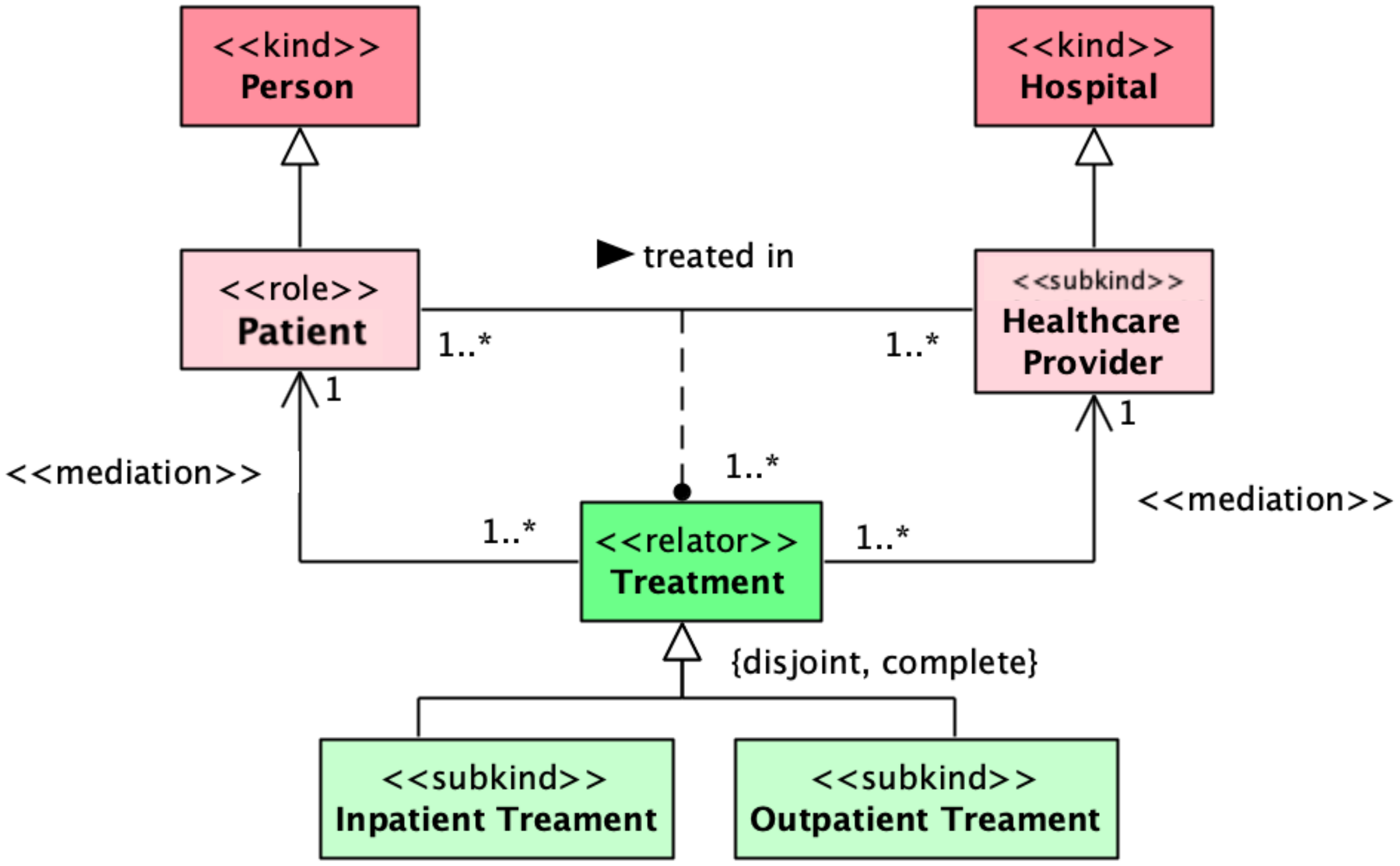

<<mediation>>
<<relator>>
Treatment
<<mediation>>
1..*
1..*
1
1..*
1..*
<<role>>
Patient
1..*
1..*
1..*
<<subkind>>
Healthcare Provider
<<material>>
treated in

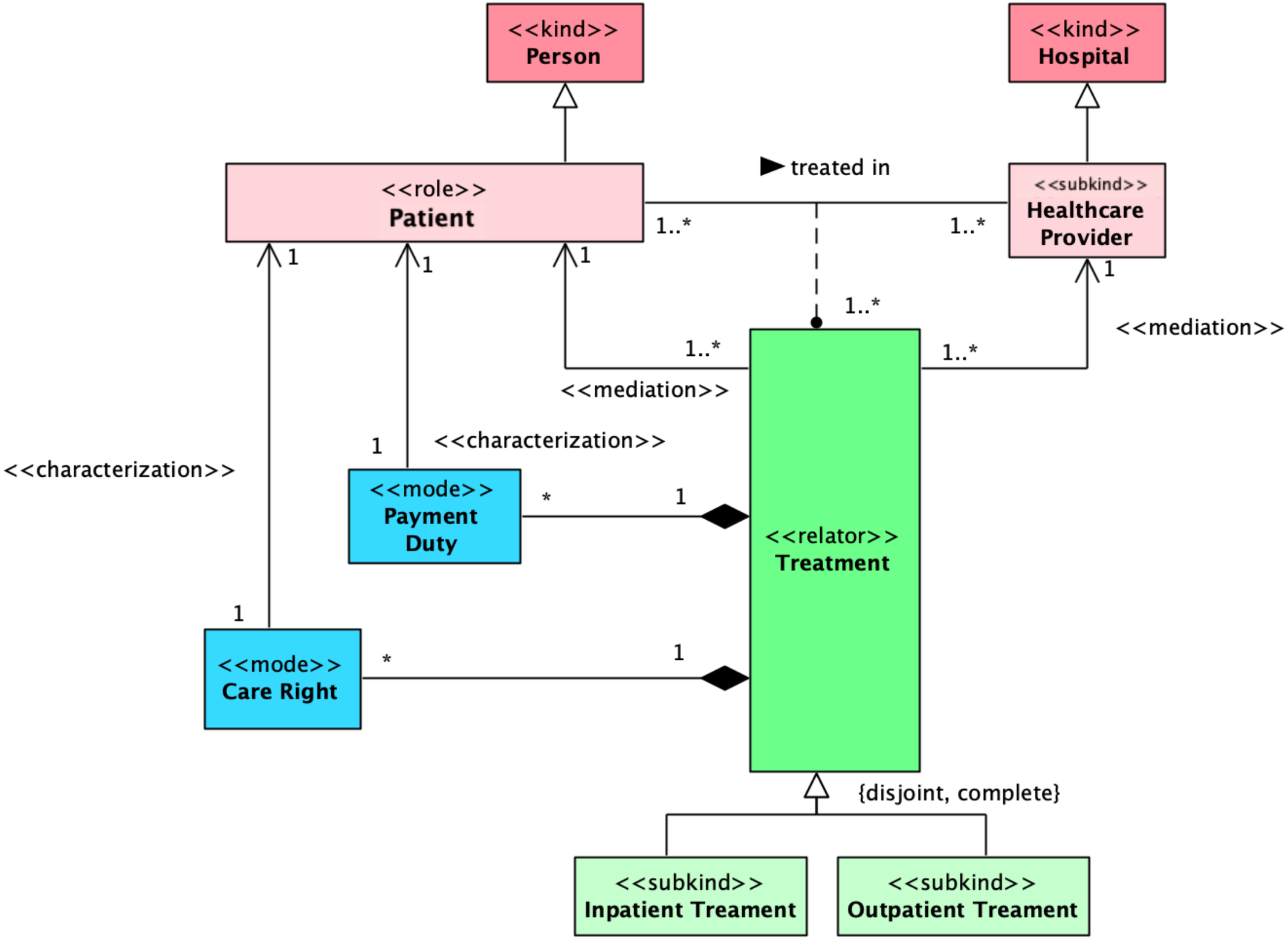

has more serious medical condition
*
*
Person
treated by
1..*
*
Health Care Provider

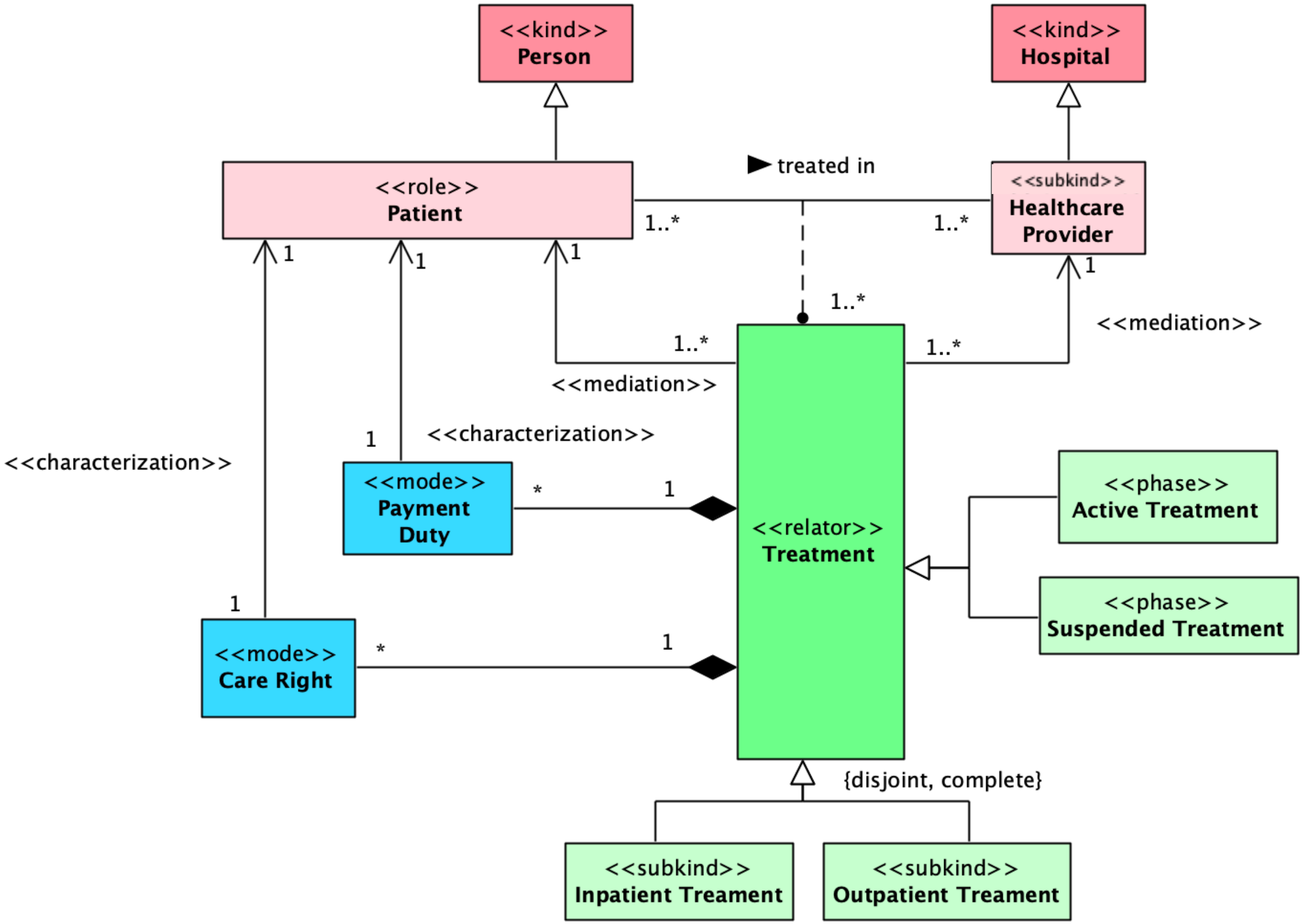

<<kind>>
Person
{disjoint, complete}
<<kind>>
Organization
<<comparative>>
has more serious medical
condition than/
<<phase>>
Healthy Person
<<phase>>
Unhealthy Person
*
*
1
1..*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder
<<role>>
Patient
1..*
1
treated by
*
<<subkind>>
Health Care Provider
1..*
1..*
1..*
<<relator>>
Treatment
1..*
<<mediation>>
<<mediation>>

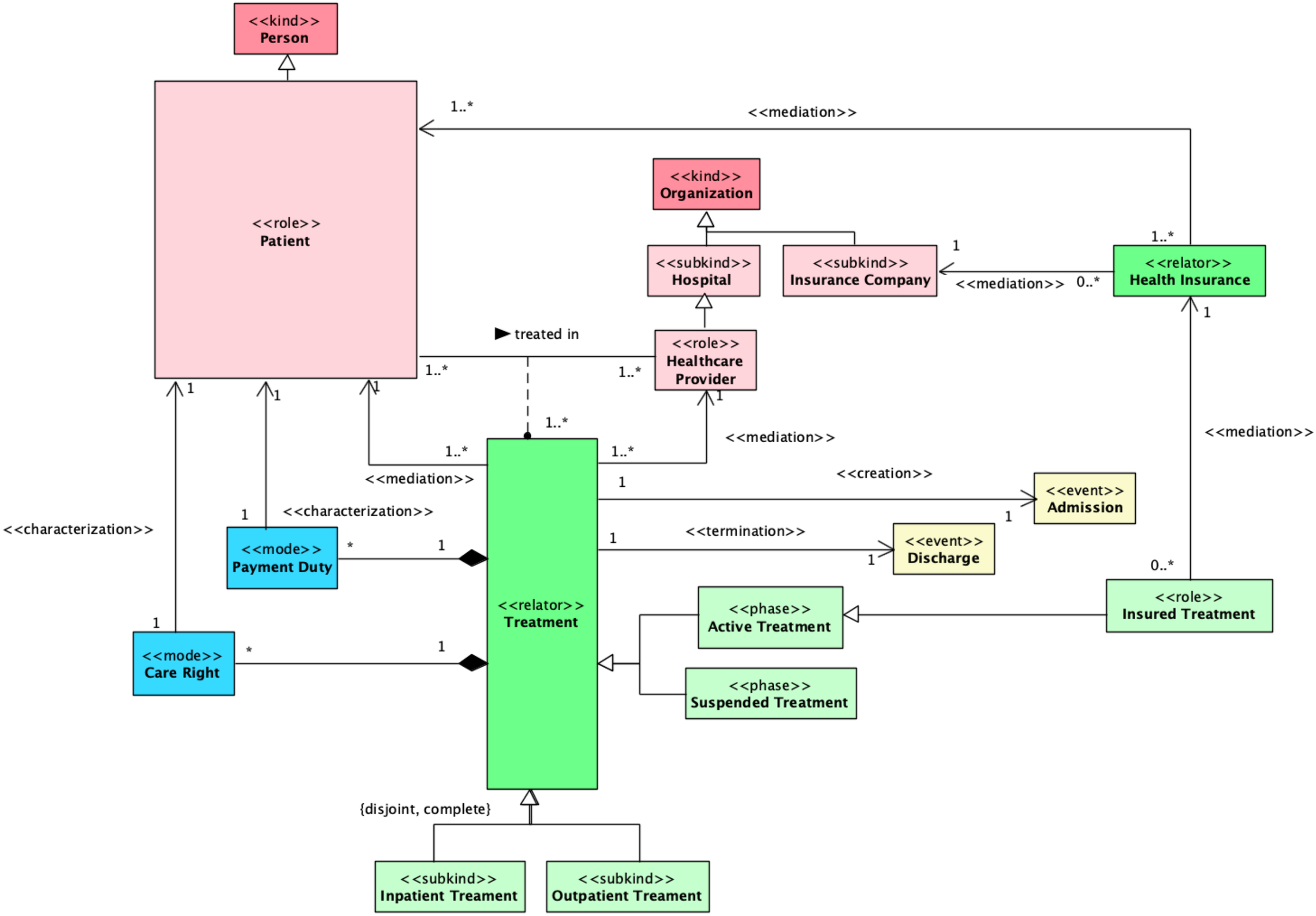
Unpacking Relations

1. Truthmaking
2. Disambiguation (**Semantic Clarity**)
3. Helps to elicit tacit knowledge that would otherwise remain tacit (**Completeness**)
4. Makes justice to the complexity of the relational phenomena (**Precision**)
5. It helps differentiating subsetting, redefinition and specialisation between relations
6. It solves the problem of transitivity of parenthood
7. It disambiguates between relators and events


<<kind>>
Person
<<kind>>
Hospital
<<role>>
Patient
treated in
<<subkind>>
Healthcare
Provider
1..*
1..*
1
1
<<mediation>>
1..*
<<mediation>>
1..*
<<relator>>
Treatment
1..*
{disjoint, complete}
<<subkind>>
Inpatient Treament
<<subkind>>
Outpatient Treament


<<kind>>
Person
<<kind>>
Hospital
<<role>>
Patient
treated in
<<subkind>>
Healthcare
Provider
1..*
1..*
1
1
1
1..*
1..*
1..*
<<mediation>>
<<mediation>>
1
<<characterization>>
<<characterization>>
<<mode>>
Payment
Duty
*
1
<<relator>>
Treatment
1
<<mode>>
Care Right
*
1
{disjoint, complete}
<<subkind>>
Inpatient Treament
<<subkind>>
Outpatient Treament


<<kind>>
Person
<<kind>>
Hospital
<<role>>
Patient
treated in
<<subkind>>
Healthcare
Provider
1..*
1..*
1
1
1
1..*
1..*
1..*
<<mediation>>
<<mediation>>
1
<<characterization>>
<<characterization>>
<<mode>>
Payment
Duty
*
1
<<relator>>
Treatment
<<phase>>
Active Treatment
<<phase>>
Suspended Treatment
1
<<mode>>
Care Right
*
1
{disjoint, complete}
<<subkind>>
Inpatient Treament
<<subkind>>
Outpatient Treament

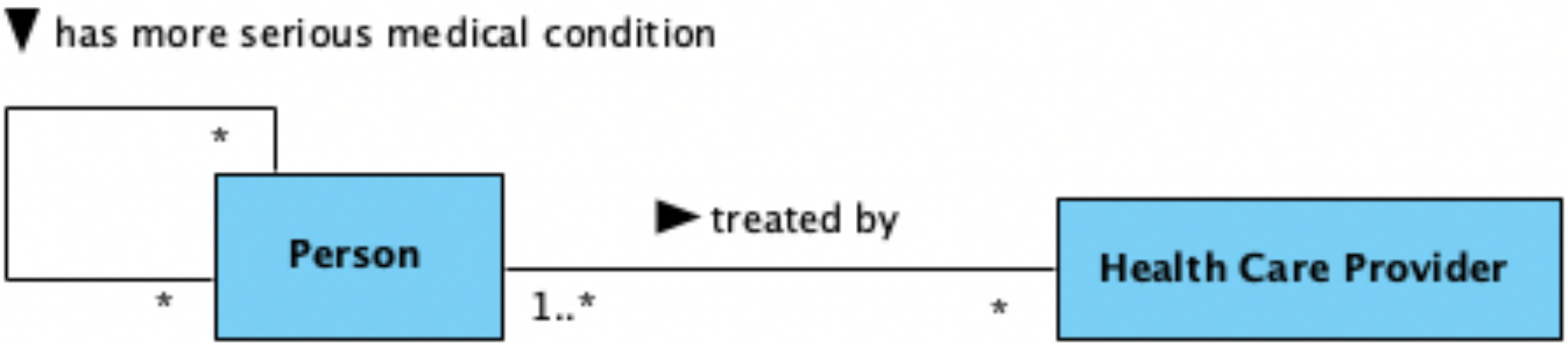

<<kind>>
Person
<<role>>
Patient
1..*
<<mediation>>
<<kind>>
Organization
<<subkind>>
Hospital
<<subkind>>
Insurance Company
1
<<mediation>>
0..*
1..*
<<relator>>
Health Insurance
1
<<mediation>>
treated in
1..*
1..*
<<role>>
Healthcare
Provider
1
1..*
<<mediation>>
1..*
1..*
<<mediation>>
<<creation>>
1
<<event>>
Admission
1
<<characterization>>
1
<<characterization>>
<<mode>>
Payment Duty
*
1
<<termination>>
1
<<event>>
Discharge
1
<<relator>>
Treatment
0..*
<<role>>
Insured Treatment
1
<<mode>>
Care Right
*
1
<<phase>>
Active Treatment
<<phase>>
Suspended Treatment
{disjoint, complete}
<<subkind>>
Inpatient Treament
<<subkind>>
Outpatient Treament

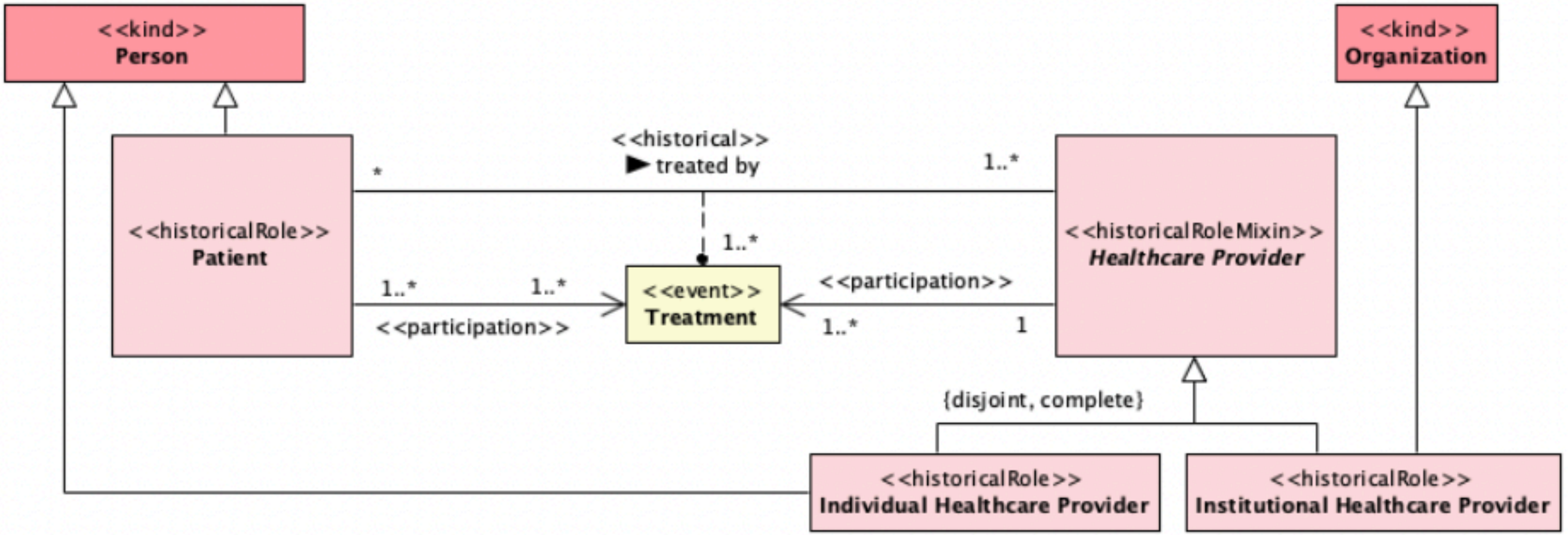
The difference between these models is not just one of **expressivity** but one of **nature**!
(Descriptive x Explanatory)

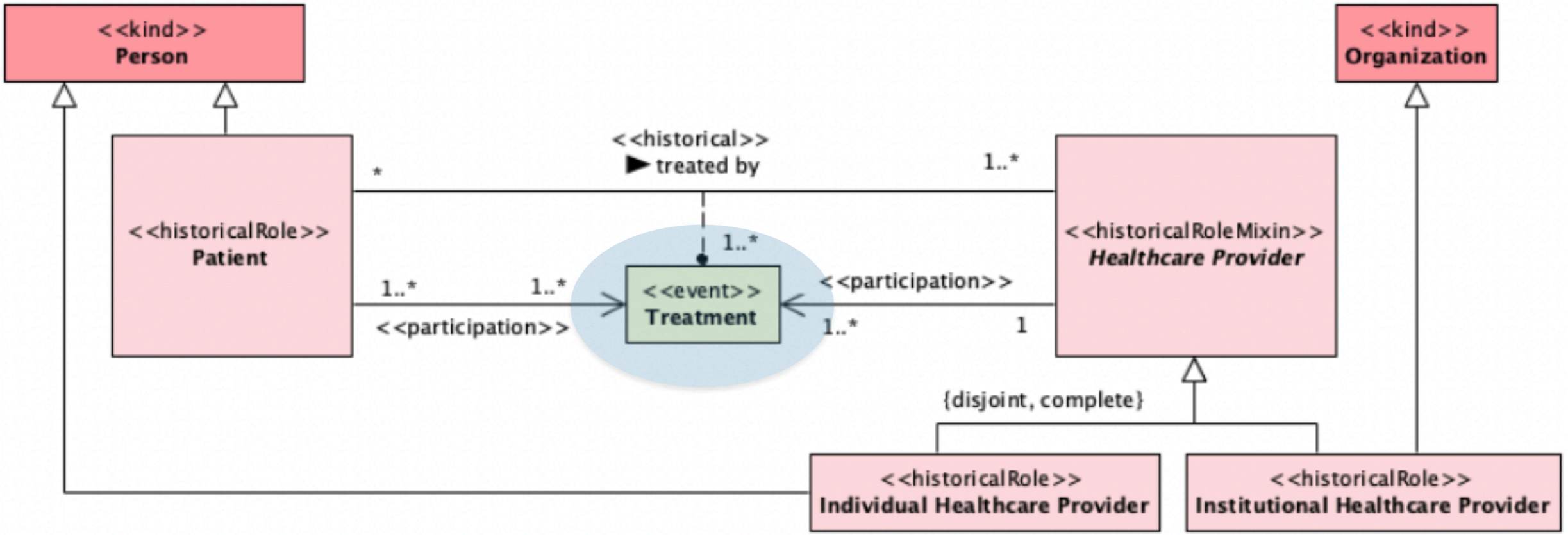
"Traditionally, theories are said to bear two sorts of relation to the observable phenomena: **description** and **explanation**. Description can be more or less accurate, more or less informative; as a minimum, the facts must 'be allowed by the theory' (fit some of its models), as a maximum the theory actually implies the facts in question."

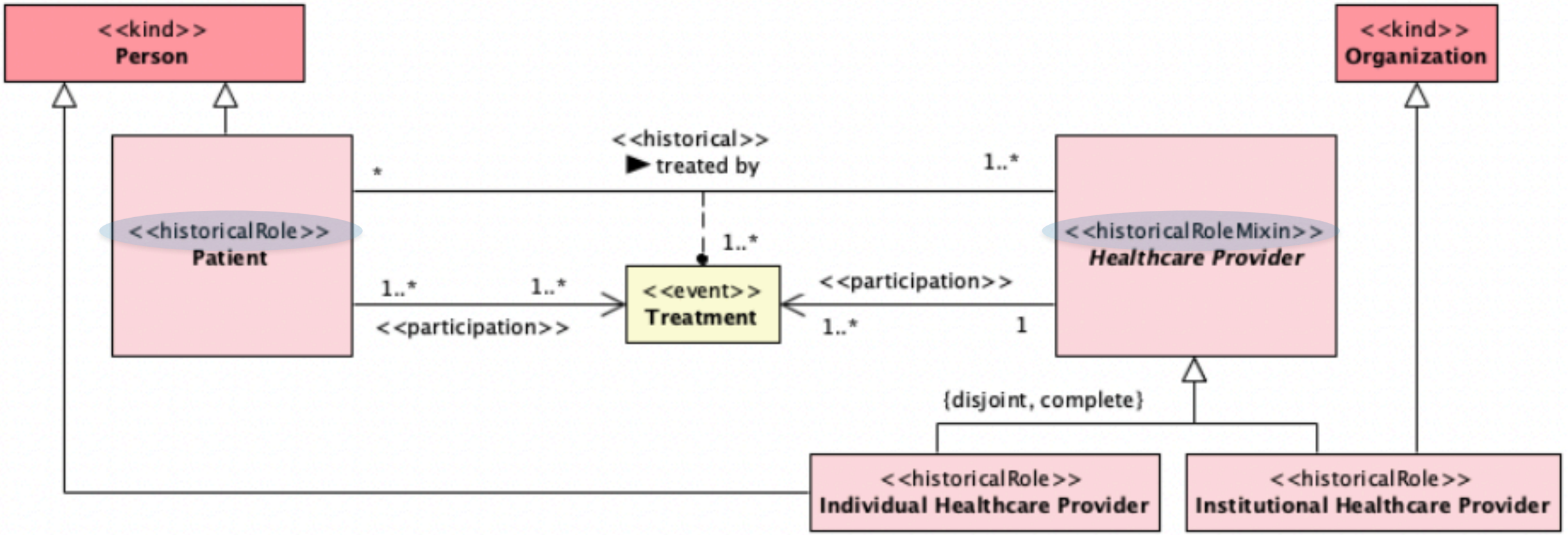
“**Explanation is…‘over and above’ description**; for example, Boyle’s law describes the relationship between the pressure, temperature, and volume of a contained gas, but does not explain it - kinetic theory explains it”
(Bas van Fraasen)

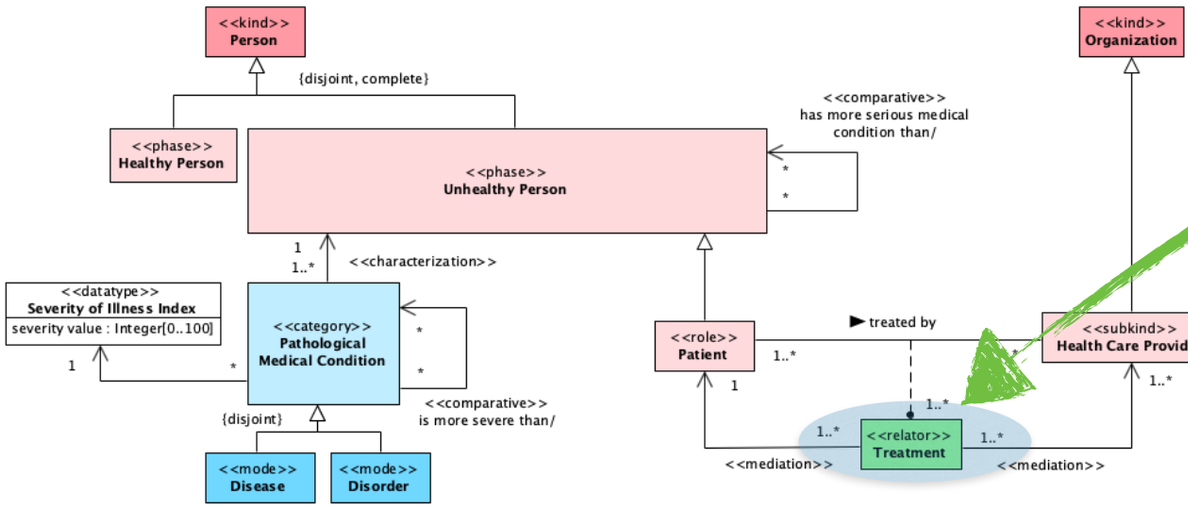
To **explain** is to **reveal** one's **ontological commitment** (**real-world semantics**) and that is what is **needed for semantic interoperability**


has more serious medical condition
*
*
Person
treated by
1..*
*
Health Care Provider

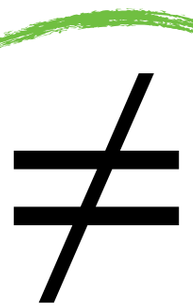

<<kind>>
Person
<<kind>>
Organization
<<historical>>
treated by
*
1..*
<<historicalRole>>
Patient
<<historicalRoleMixin>>
Healthcare Provider
1..*
1..*
1..*
<<participation>>
<<event>>
Treatment
<<participation>>
1..*
1
{disjoint, complete}
<<historicalRole>>
Individual Healthcare Provider
<<historicalRole>>
Institutional Healthcare Provider

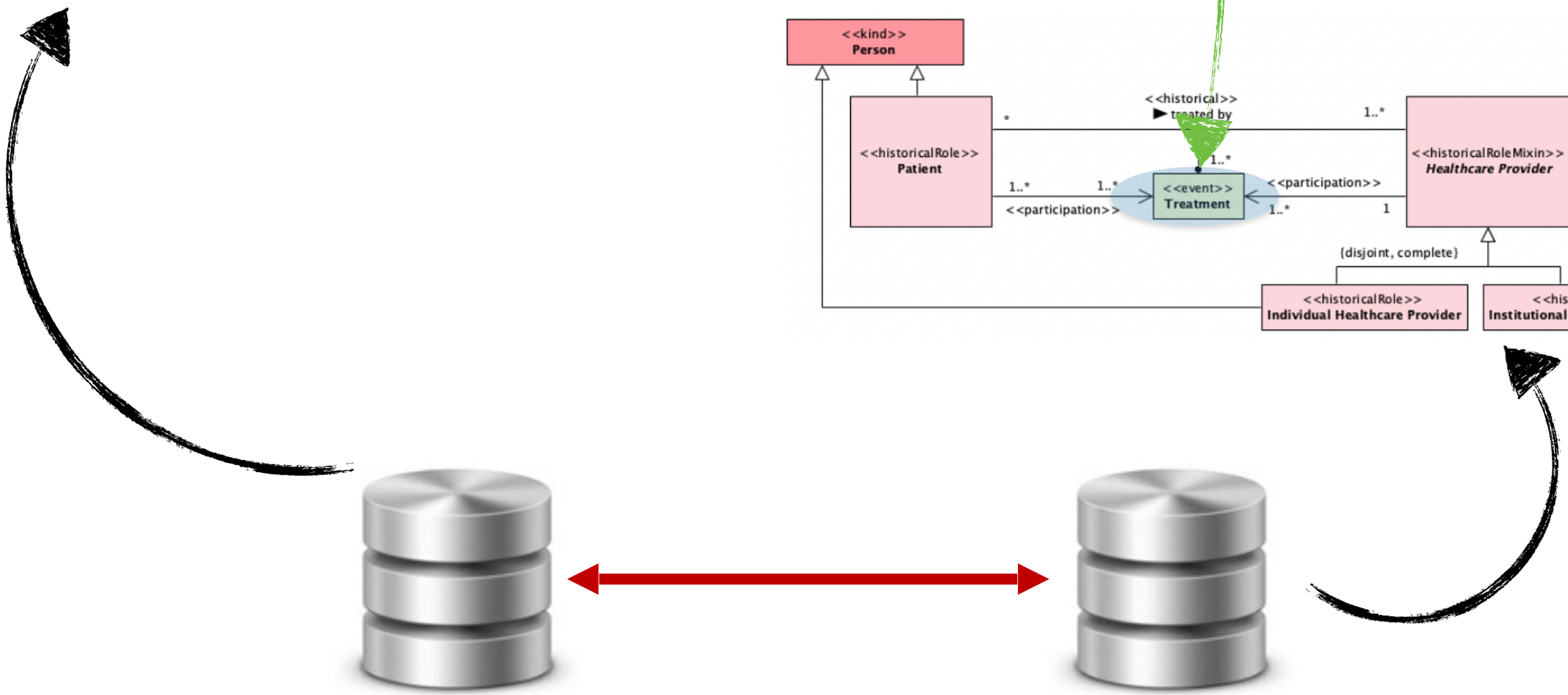

<<kind>>
Person
<<kind>>
Organization
<<historical>>
treated by
*
1..*
<<historicalRole>>
Patient
<<historicalRoleMixin>>
Healthcare Provider
1..*
1..*
1..*
<<participation>>
<<event>>
Treatment
<<participation>>
1..*
1
{disjoint, complete}
<<historicalRole>>
Individual Healthcare Provider
<<historicalRole>>
Institutional Healthcare Provider

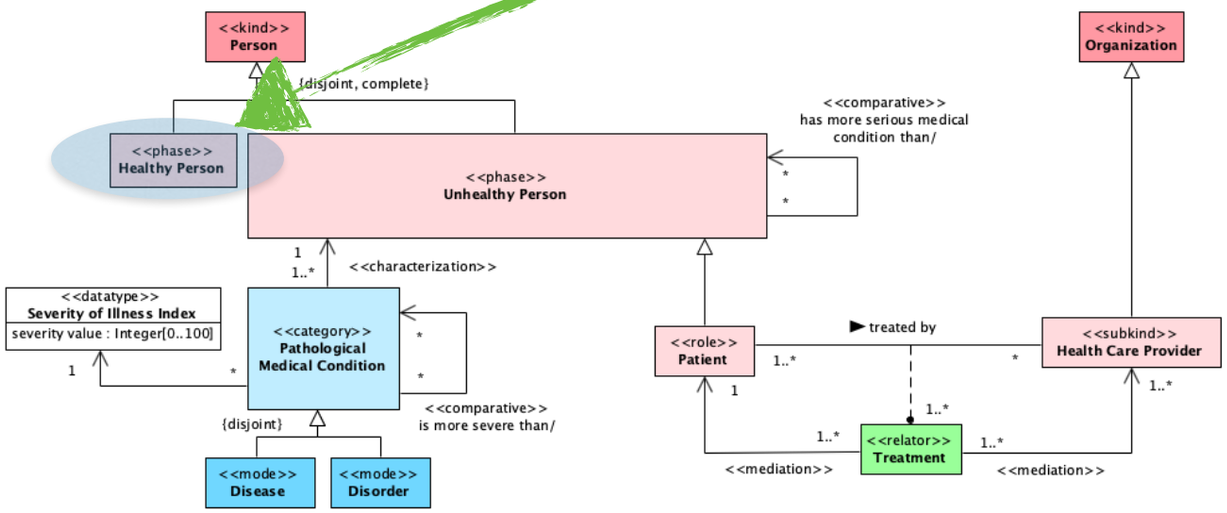

<<kind>>
Person
<<kind>>
Organization
<<historical>>
treated by
*
1..*
1..*
<<historicalRole>>
Patient
<<historicalRoleMixin>>
Healthcare Provider
1..*
1..*
<<participation>>
<<event>>
Treatment
<<participation>>
1..*
1
{disjoint, complete}
<<historicalRole>>
Individual Healthcare Provider
<<historicalRole>>
Institutional Healthcare Provider


<<kind>>
Person
{disjoint, complete}
<<phase>>
Healthy Person
<<phase>>
Unhealthy Person
<<comparative>>
has more serious medical
condition than/
1
1..*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder

<<kind>>
Organization
<<role>>
Patient
1..*
1
treated by
<<subkind>>
Health Care Provider
1..*
1..*
1..*
1..*
<<relator>>
Treatment
1..*
<<mediation>>
<<mediation>>

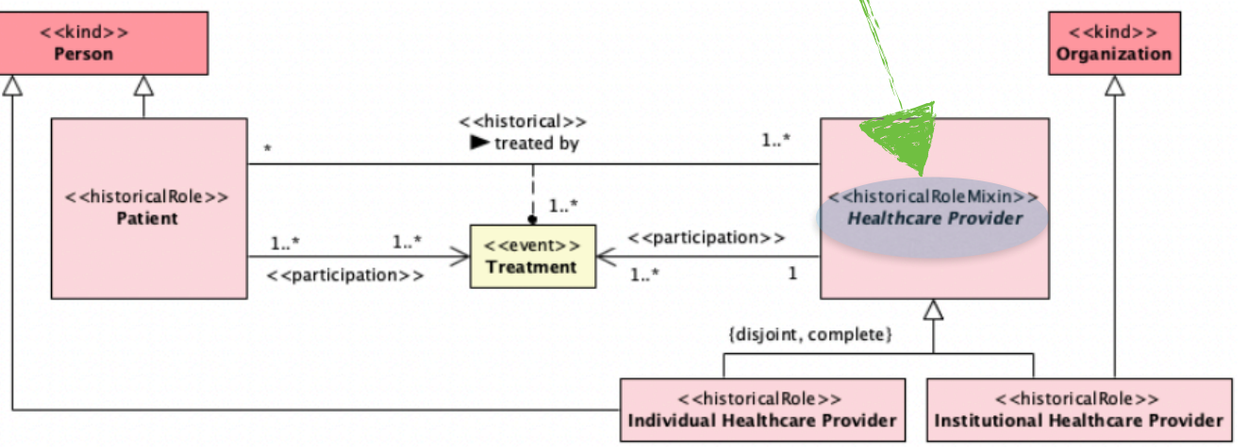



<<kind>>
Person
<<historicalRole>>
Patient
*
<<historical>>
treated by
1..*
1..*
1..*
1..*
<<event>>
Treatment
<<participation>>
<<participation>>
1..*
1
<<historicalRoleMixin>>
Healthcare Provider
{disjoint, complete}
<<historicalRole>>
Individual Healthcare Provider
<<historicalRole>>
Institutional Healthcare Provider
<<kind>>
Organization


<<kind>>
Person
{disjoint, complete}
<<phase>>
Healthy Person
<<phase>>
Unhealthy Person
<<comparative>>
has more serious medical
condition than/
1
1..*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder

<<kind>>
Organization
<<role>>
Patient
1
treated by
1..*
*
<<subkind>>
Health Care Provider
1..*
1..*
<<relator>>
Treatment
1..*
1..*
<<mediation>>
<<mediation>>

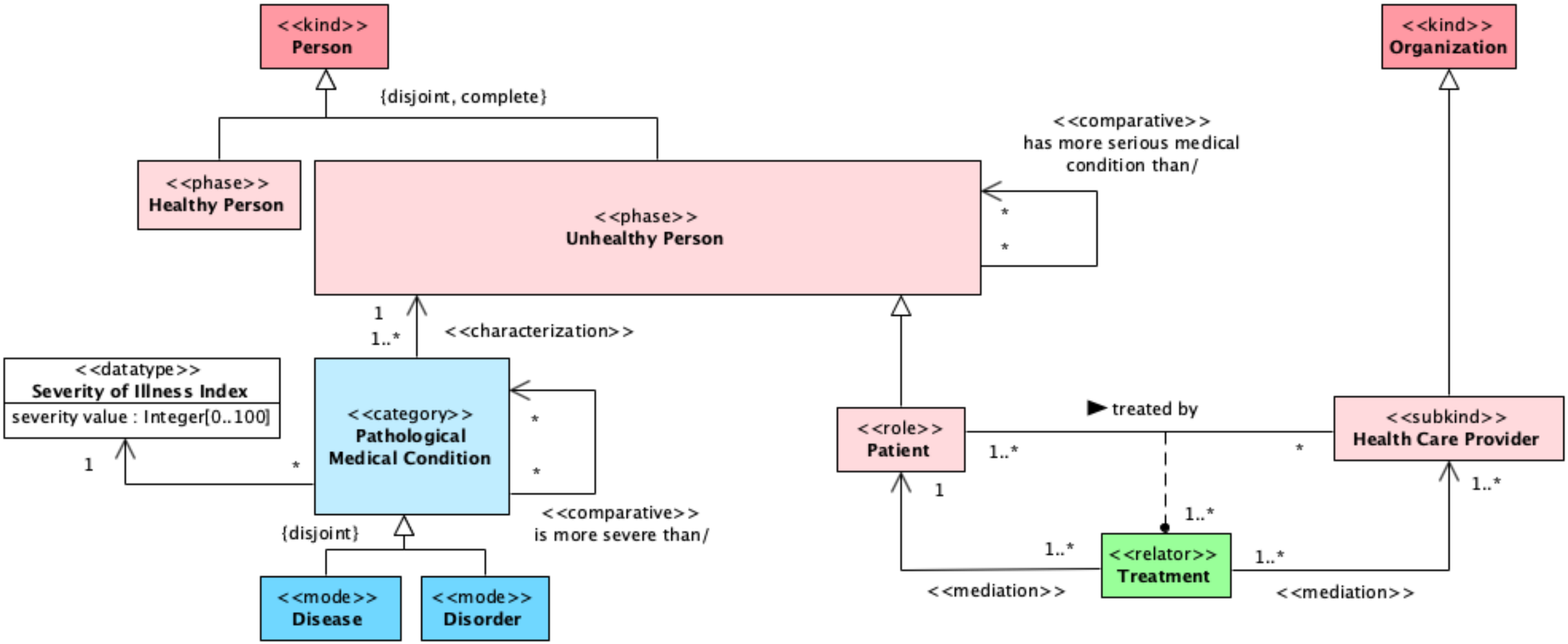
≠


<<kind>>
Person
<<kind>>
Organization
<<historical>>
treated by
*
1..*
<<historicalRole>>
Patient
1..*
1..*
<<participation>>
<<event>>
Treatment
<<participation>>
1..*
1
<<historicalRoleMixin>>
Healthcare Provider
{disjoint, complete}
<<historicalRole>>
Individual Healthcare Provider
<<historicalRole>>
Institutional Healthcare Provider

1. **Primitives** reflecting ontological **distinctions**

2. **Grammar** reflecting ontological **axiomatization**

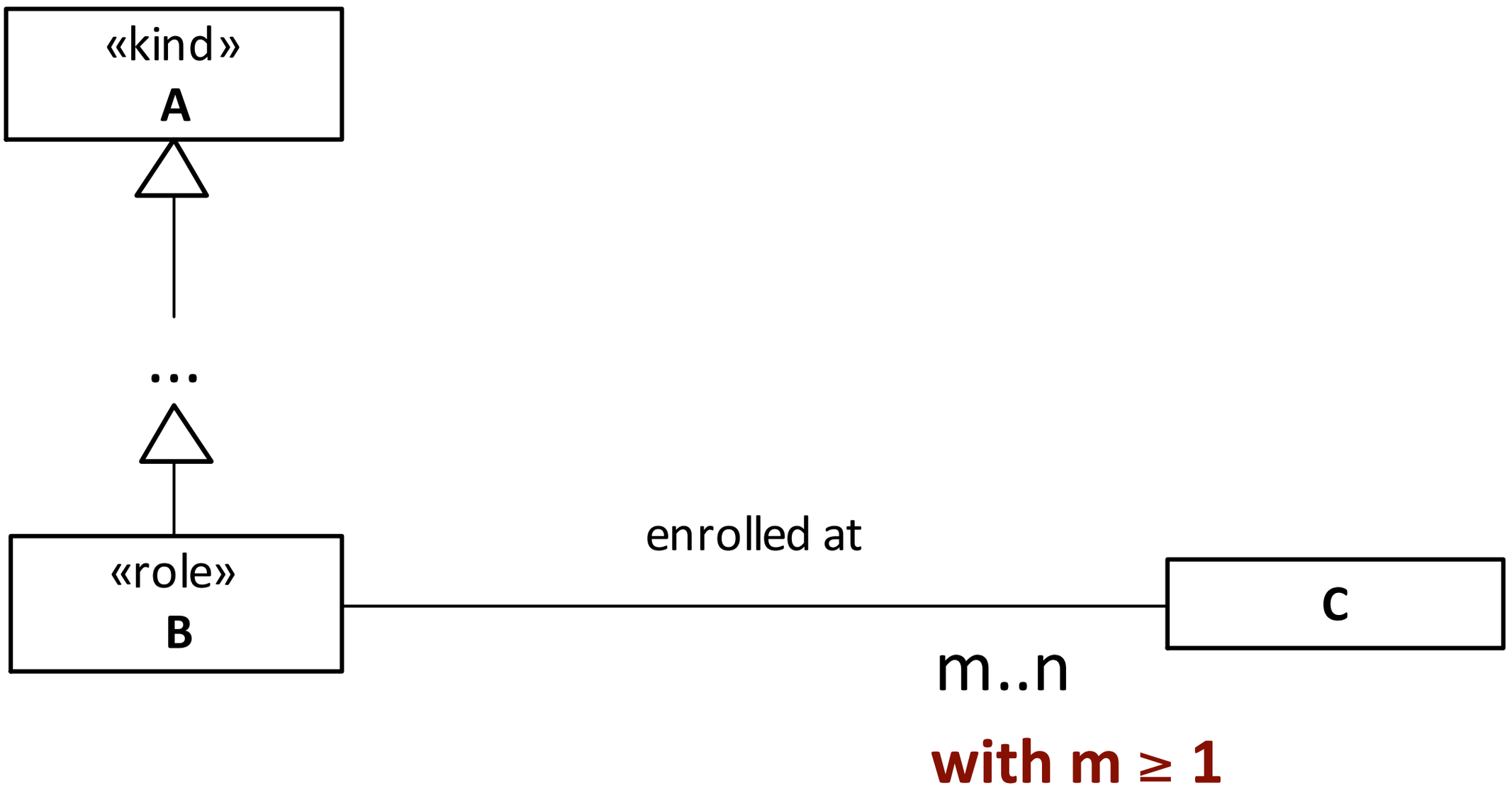
3. **Patterns** reflecting ontological **micro-theories**

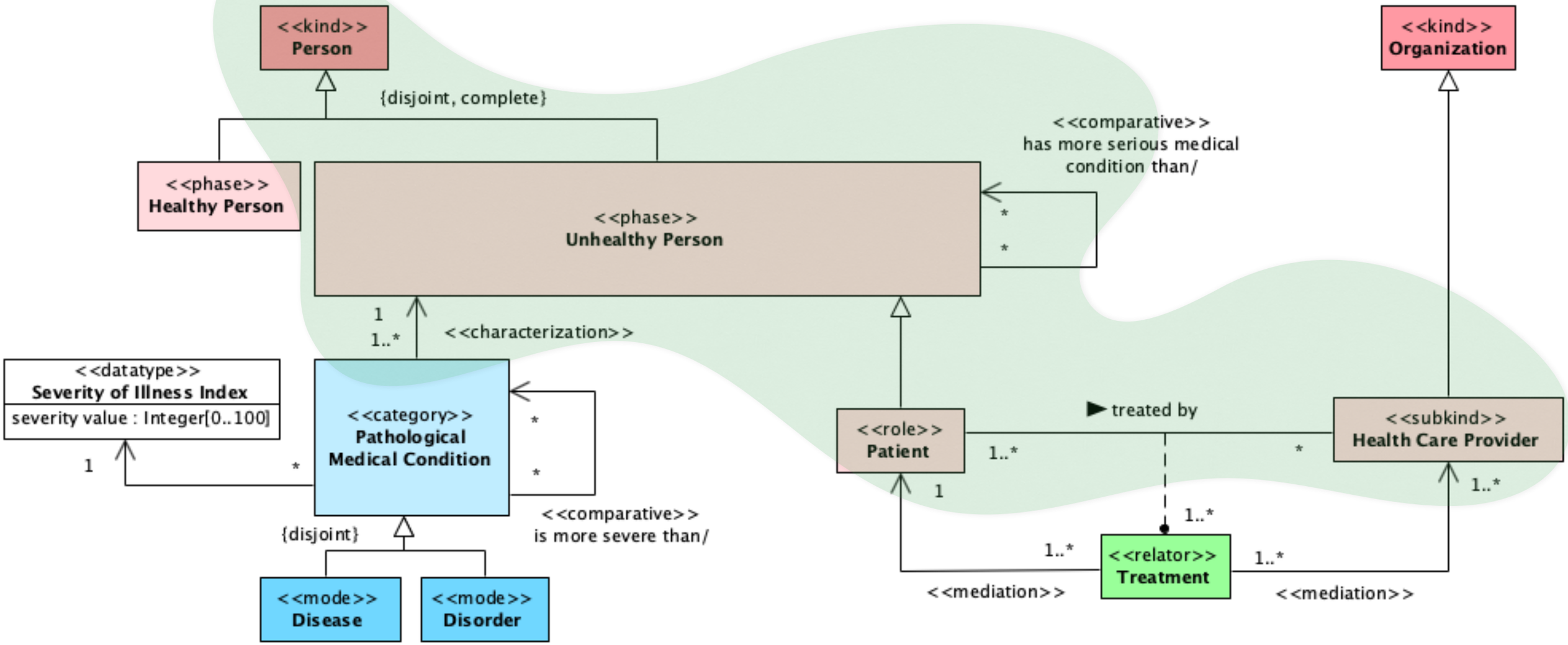

<<kind>>
Person
{disjoint, complete}
<<kind>>
Organization
<<comparative>>
has more serious medical
condition than/
<<phase>>
Healthy Person
<<phase>>
Unhealthy Person
*
*
1
1..*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder
<<role>>
Patient
1..*
1
treated by
*
<<subkind>>
Health Care Provider
1..*
1..*
1..*
<<relator>>
Treatment
1..*
<<mediation>>
<<mediation>>

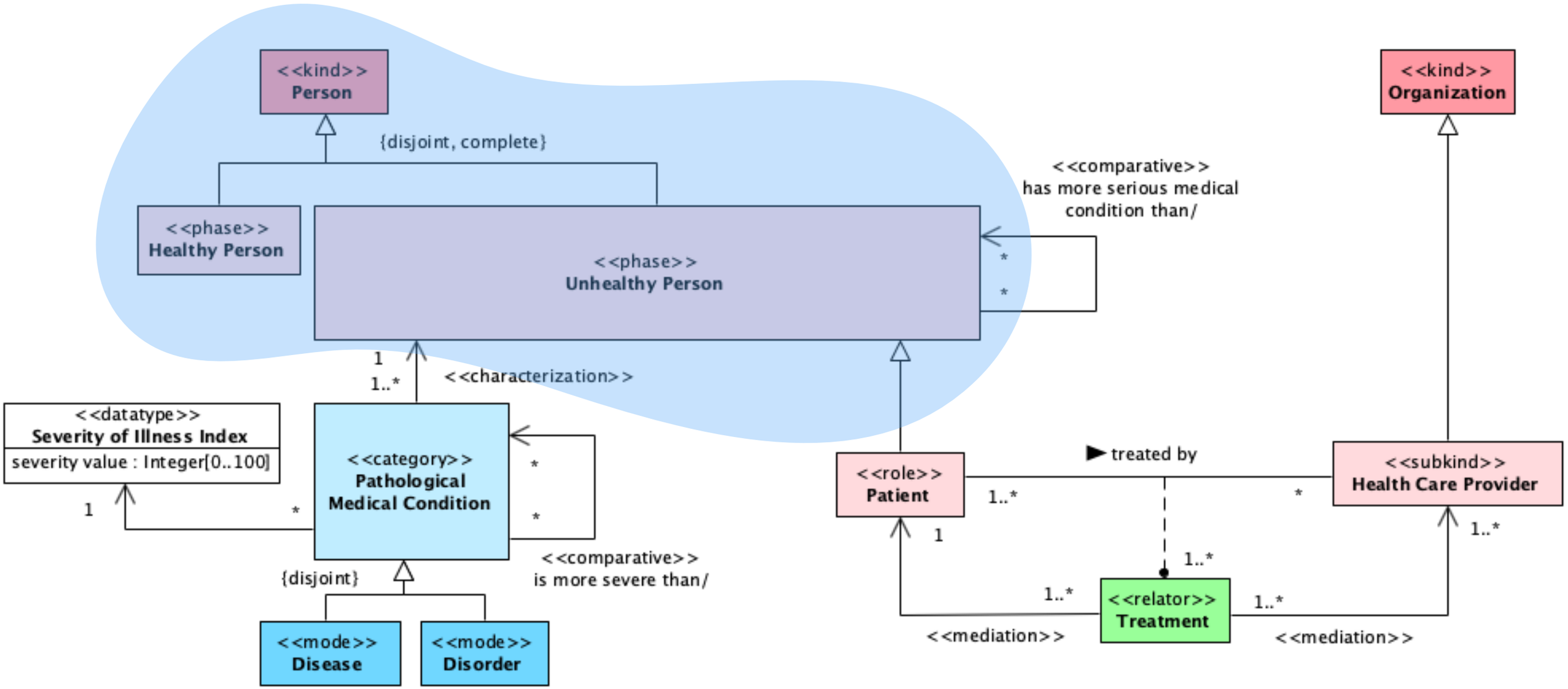
Role

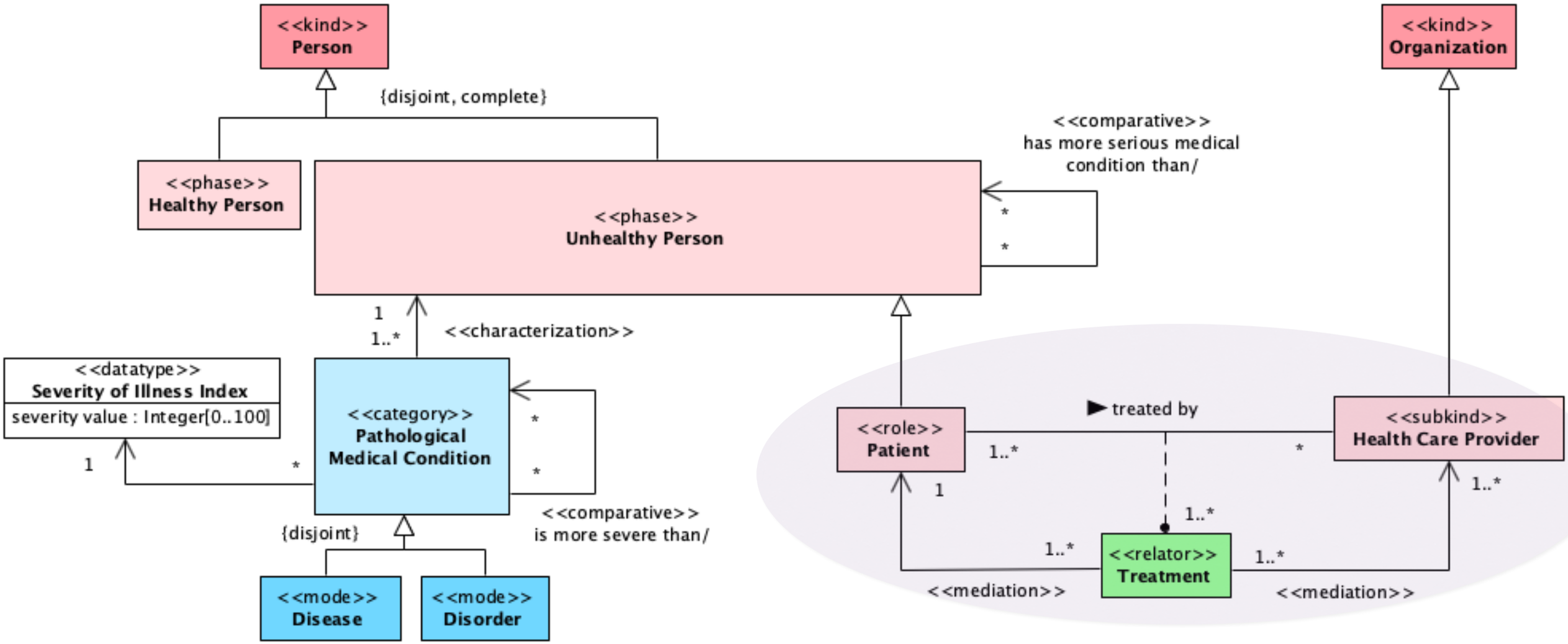
1. All instances of a given ROLE are of the **same KIND** (e.g., all Students are Person)
2. All instances of a ROLE instantiate that type only **contingently** (e.g., no Student is necessarily a Student)
3. Instances of a KIND instantiate that ROLE when participating in a certain **relational context** (e.g., instances of Person instantiate the Role Student when enrolled in na Educational Institution)
4. A ROLE cannot be a supertype of a KIND

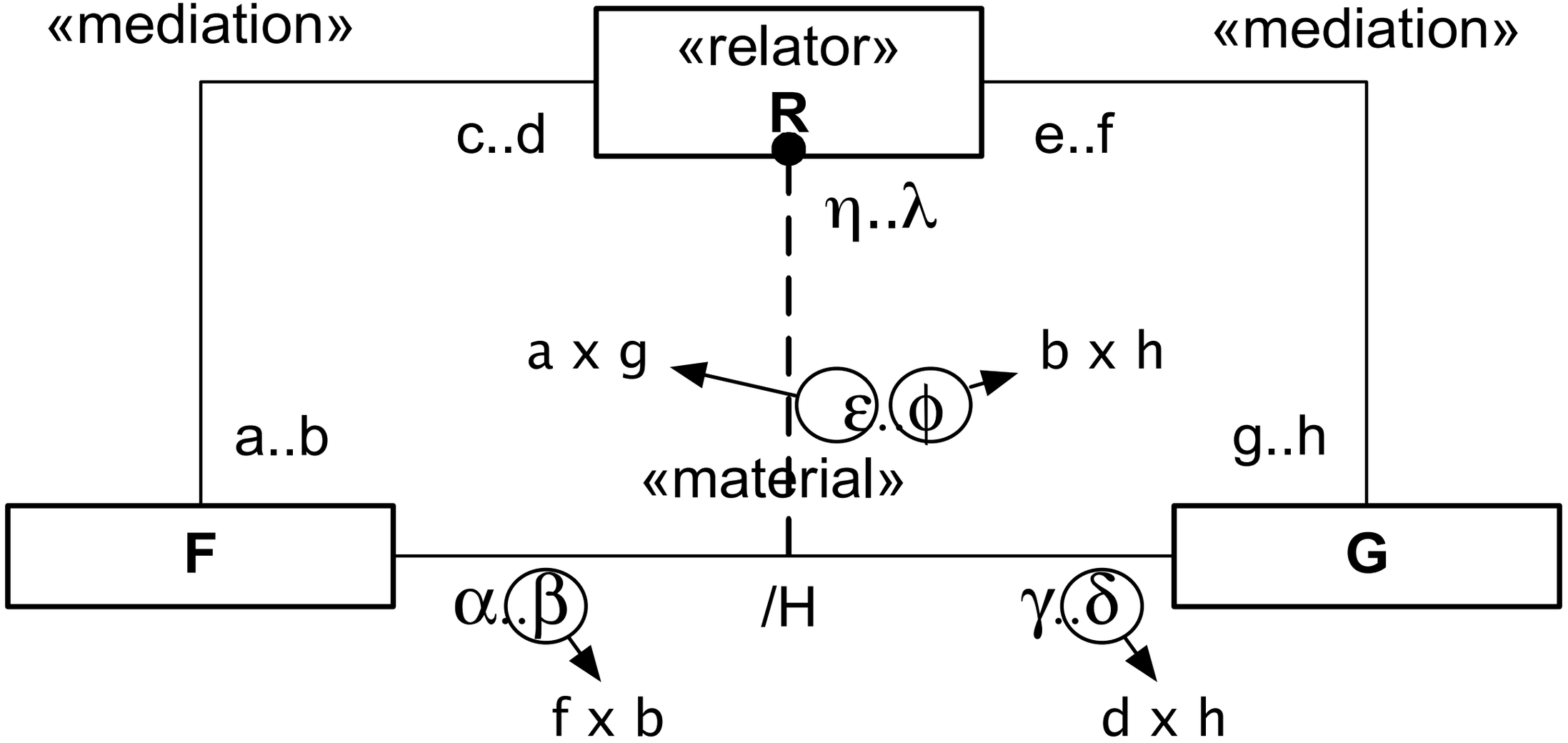
The Emerging **Role** Pattern




<<kind>>
Person
{disjoint, complete}
<<kind>>
Organization
<<phase>>
Healthy Person
<<phase>>
Unhealthy Person
<<comparative>>
has more serious medical
condition than/
*
*
1
1..*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder
treated by
<<role>>
Patient
1..*
*
<<subkind>>
Health Care Provider
1
1..*
1..*
1..*
<<relator>>
Treatment
1..*
<<mediation>>
<<mediation>>

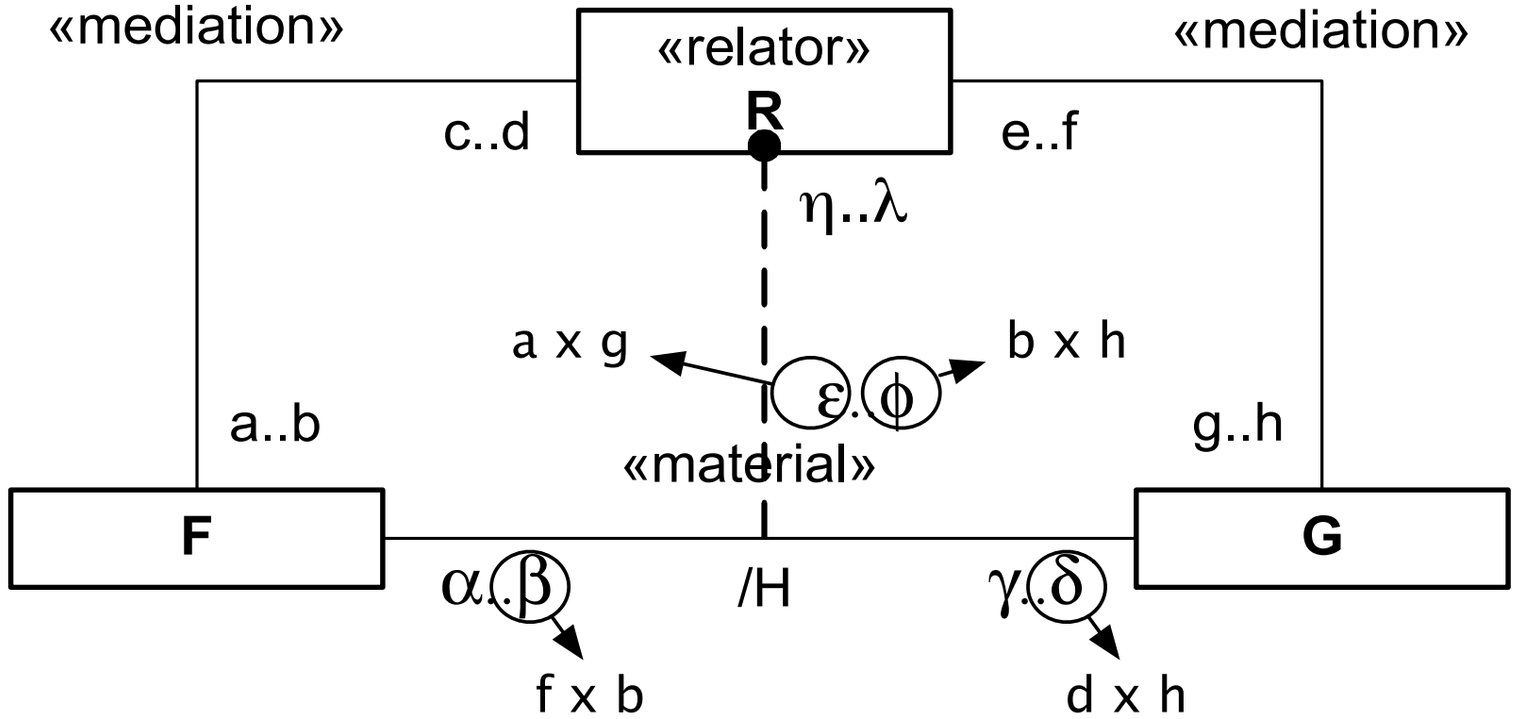

<<kind>>
Person
{disjoint, complete}
<<phase>>
Healthy Person
<<phase>>
Unhealthy Person
<<comparative>>
has more serious medical
condition than/
*
*
1
1..*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder
<<kind>>
Organization
<<role>>
Patient
1..*
1
treated by
*
<<subkind>>
Health Care Provider
1..*
1..*
1..*
<<relator>>
Treatment
1..*
<<mediation>>
<<mediation>>


<<kind>>
Person
{disjoint, complete}
<<kind>>
Organization
<<phase>>
Healthy Person
<<phase>>
Unhealthy Person
<<comparative>>
has more serious medical
condition than/
*
*
1
1..*
<<characterization>>
<<datatype>>
Severity of Illness Index
severity value : Integer[0..100]
1
*
<<category>>
Pathological
Medical Condition
*
*
<<comparative>>
is more severe than/
{disjoint}
<<mode>>
Disease
<<mode>>
Disorder
treated by
<<role>>
Patient
1..*
*
<<subkind>>
Health Care Provider
1
1..*
1..*
1..*
<<relator>>
Treatment
1..*
<<mediation>>
<<mediation>>

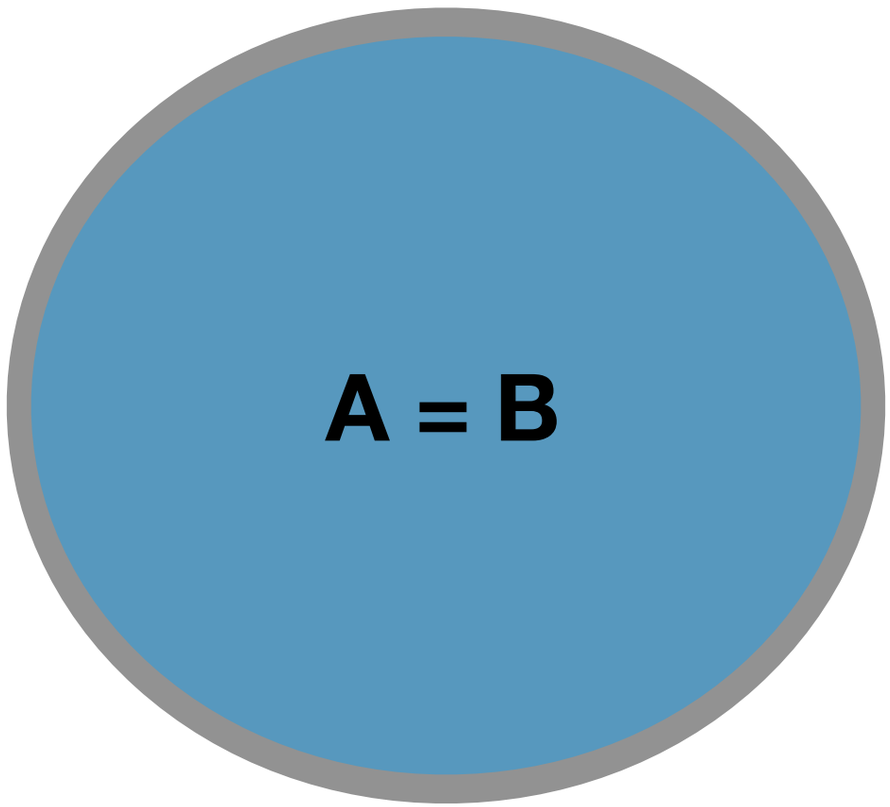


$$\forall x : F, y : G\,(H(x, y) \iff \exists r : R\,(mediation(r, x) \wedge mediation(r, y))$$

Unificatory Approach

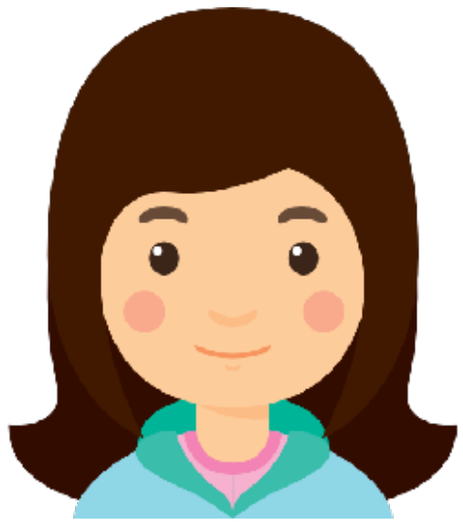


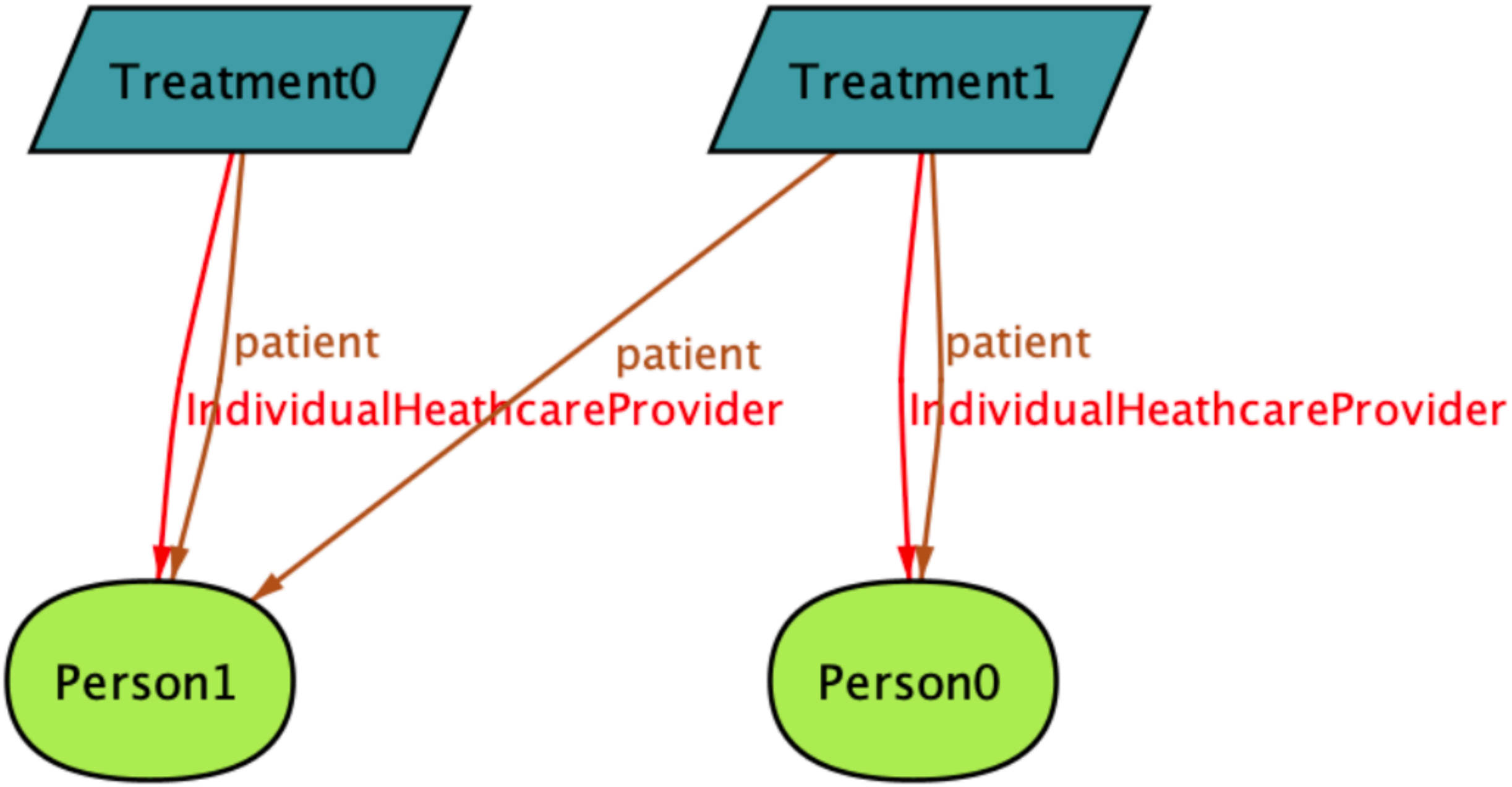
"Science advances our understanding of nature by showing us how to **derive descriptions** of many phenomena, using the same **patterns of derivation** again and again" (P. Kitcher)


A = B

A

B




Treatment0
Treatment1
patient
IndividualHeathcareProvider
patient
patient
IndividualHeathcareProvider
Person1
Person0

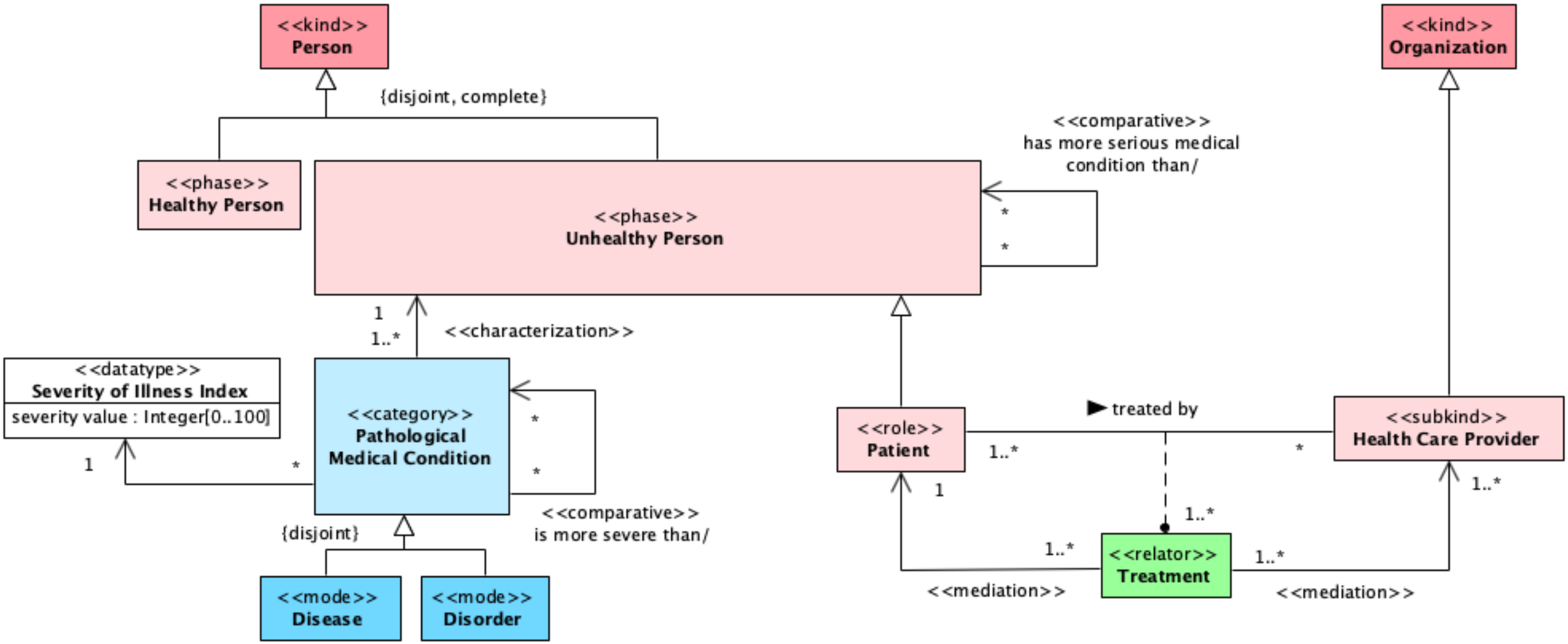
Pragmatic Explanation

1. **Requests for Explanation**, i.e., to explain is to satisfy information seeking goals of an **explanation seeker** (complexity management)

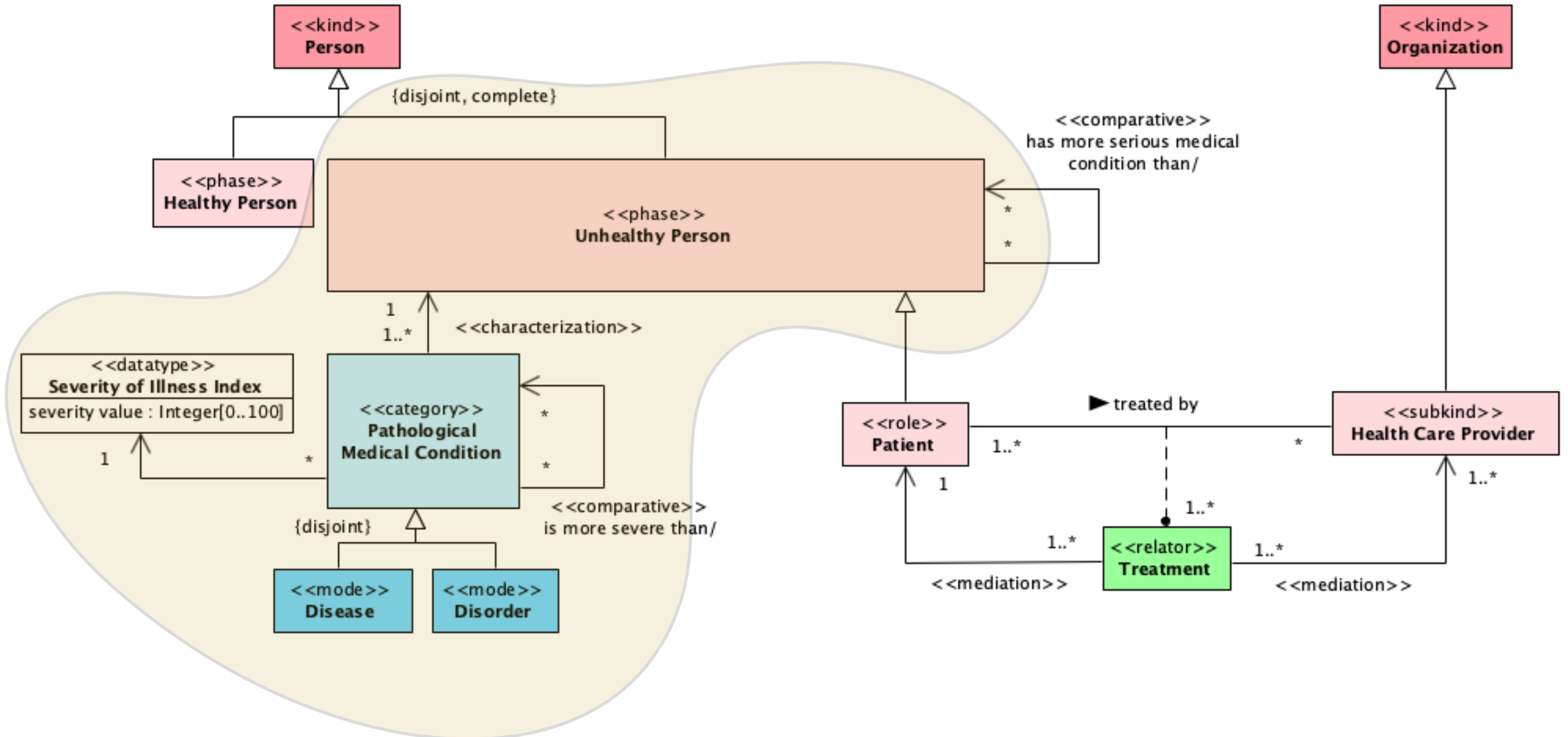
2. **Competence Questions** as Requests for Explanations

3. **Contrastive** Questions

Why is a person treated by a given healthcare provider?

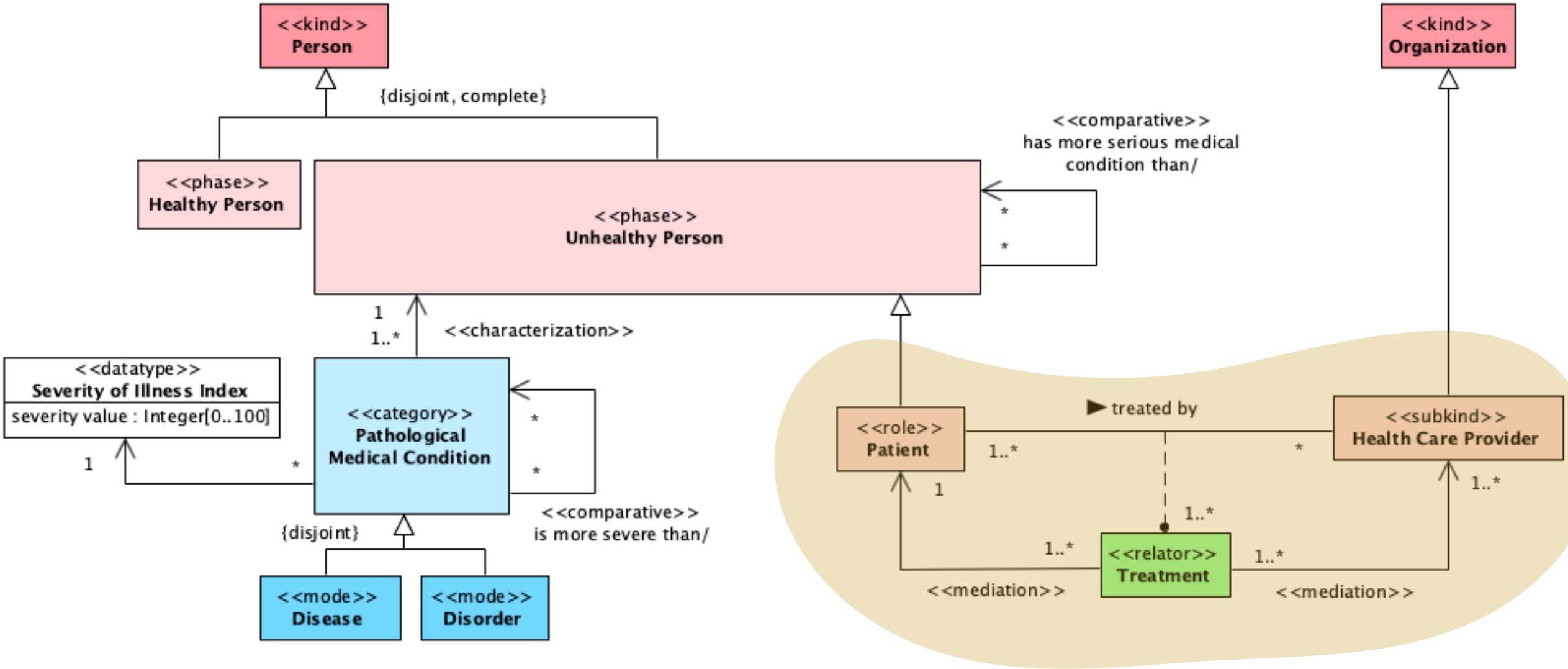


Why is a person treated by a given healthcare provider?



…**as opposed to** not being treated

Why is a person treated by a given healthcare provider?



…**as opposed to** being treated by a different healthcare provider

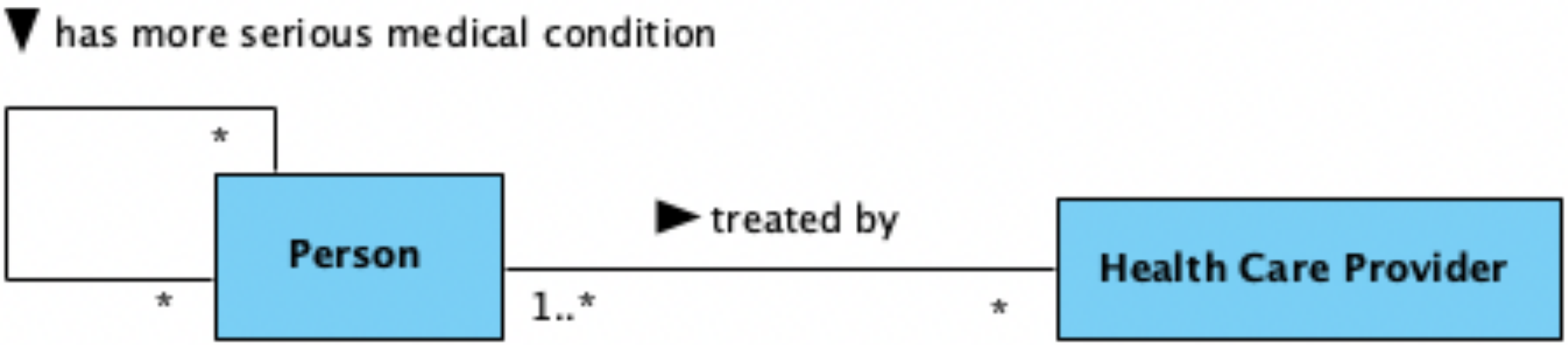
3

XAI

Explainable AI

1. Interpretability Framework or Complete Model View (“**Inherently Interpretable Models**”)

Symbolic Artifacts are **not Self-Explanatory** just in virtue of being symbolic!


has more serious medical condition
*
*
Person
treated by
1..*
*
Health Care Provider

Explainable AI

1. Interpretability Framework or Complete Model Approach (**“Inherently Interpretable Models”**)
2. Explainability Framework Partial-Model Approach

Counterfactual Explanations

"You were denied a loan because your annual income was £30,000. If your income had been £45,000, you would have been offered a loan."

Semantics and explanation: why counterfactual explanations produce adversarial examples in deep neural networks

Kieran Browne*
Research School of Humanities & the Arts
Australian National University
kieran.browne@anu.edu.au

Ben Swift
Research School of Computer Science
Australian National University
ben.swift@anu.edu.au

"…[consist] only of **semantically dense** and contextually relevant dimensions in the network's feature space…we would need to be able to **reveal the semantics** of hidden network units ('hidden neurons')… **there can be no explanation without semantics**"

“**representations in NNs are not really ‘signs’** that correspond to anything interpretable — but are distributed, correlative and continuous numeric values …a hidden unit cannot on its own represent any object that is **metaphysically meaningful**”
(Walid Saba)

Counterfactual Explanations

"You were denied a **loan** because your **annual income** was £30,000. If your income had been £45,000, you would have been **offered** a loan."

“What good is an explanation?” (Peter **Lipton**)

1. Knowing-that x Knowing-why

2. Why-Regress

3. Self-Evidencing

4. Counterfactuality/Constrastivity and **Contestability (Pearl)**

“Right to Explanation” (Z.C. **Lipton**)

1. present clear reasoning based on **falsifiable propositions**
2. offer some natural way of **contesting** these propositions and modifying the decisions appropriately if they are falsified

meaning 1 of 2 noun

mean·ing ˈmē-niŋ

Synonyms of *meaning* ›

1 **a** : the thing one intends to convey especially by language

> Do not mistake my *meaning*.

b : the thing that is conveyed especially by language

> Many words have more than one *meaning*.

2 : significant quality

especially : implication of a hidden or special significance

> a glance full of *meaning*

meaning 1 of 2 noun

mean·ing ˈmē-niŋ

Synonyms of *meaning* ›

Semantics

1 **a** : the thing one intends to convey especially by language

> Do not mistake my *meaning*.

b : the thing that is conveyed especially by language

> Many words have more than one *meaning*.

2 : significant quality

especially : implication of a hidden or special significance

> a glance full of *meaning*

meaning 1 of 2 noun

mean·ing ˈmē-niŋ

Synonyms of *meaning* ›

Goal-Based

1 **a** : the thing one intends to convey especially by language

> Do not mistake my *meaning*.

b : the thing that is conveyed especially by language

> Many words have more than one *meaning*.

2 : significant quality

especially : implication of a hidden or special significance

> a glance full of *meaning*

Final **Take Away** Messages

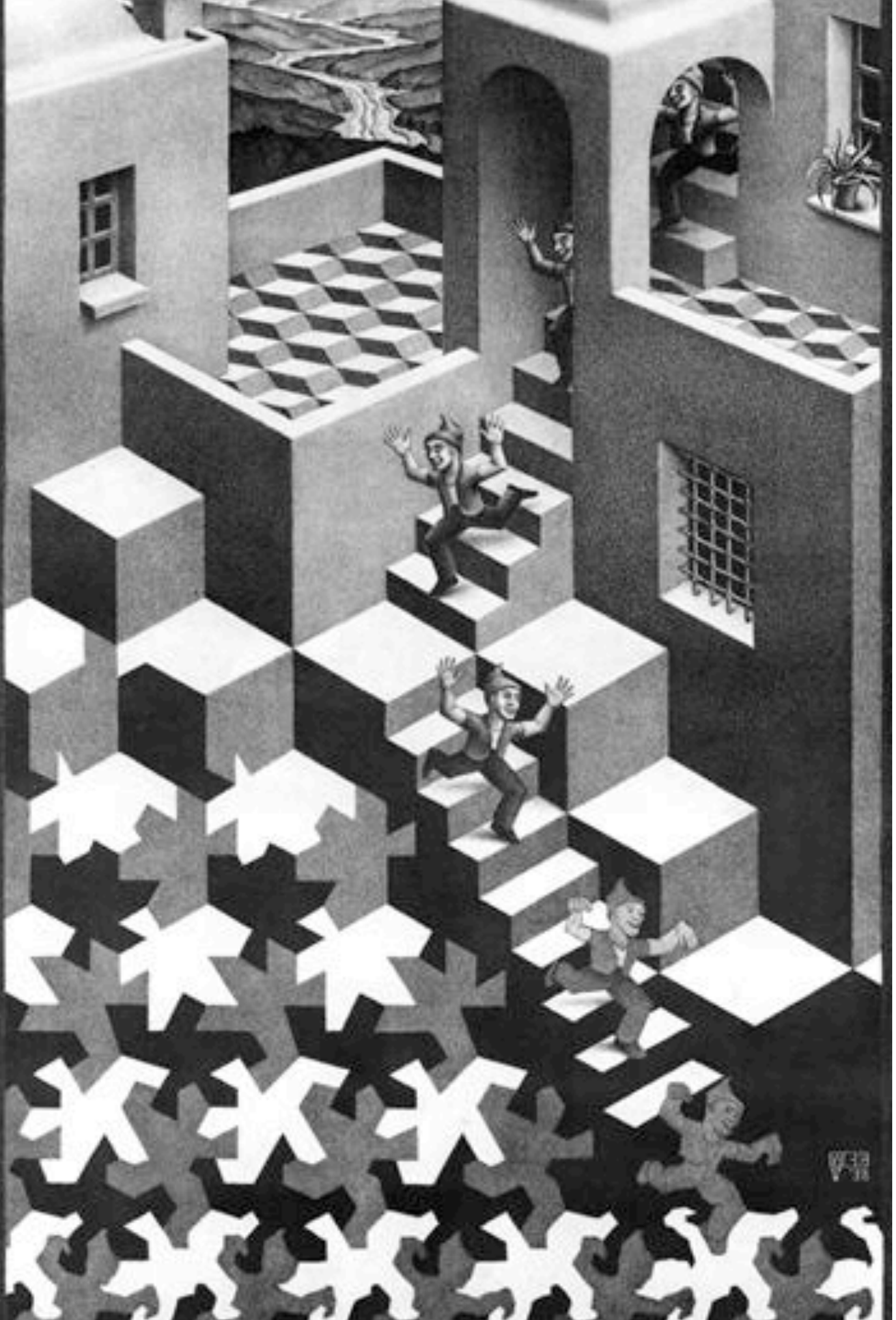
1. “No **Explanation** without **Semantics**”

Final **Take Away** Messages

1. “No **Explanation** without **Semantics**”
2. No **Semantics** without **ontology**

Final **Take Away** Messages

1. “No **Explanation** without **Semantics**”
2. No **Semantics** without **ontology**
3. No **ontology** without **Ontology**



UNIVERSITEIT
TWENTE.

SCS