From Data to Knowledge through Grailog Visualization

(Long version: http://www.cs.unb.ca/~boley/talks/RuleMLGrailog.pdf)

Harold Boley Faculty of Computer Science University of New Brunswick Fredericton, NB, Canada

ISO 15926 and Semantic Technologies 2013 Conference Sogndal, Norway, 5-6 September 2013

Acknowledgements

Thanks for feedback on various versions and parts of this presentation (the long version has all parts, hence gapless slide numbers):

Grailog 1.0: Graph-Logic Visualization of Ontologies and Rules

The 7th International Web Rule Symposium (RuleML 2013), University of Washington, Seattle WA, 11-13 July 2013

The Grailog Systematics for Visual-Logic Knowledge Representation with Generalized Graphs Faculty of Computer Science Seminar Series, University of New Brunswick, Fredericton, Canada, 26 September 2012 High Performance Computing Center Stuttgart (HLRS), Stuttgart, Germany, 14 August 2012

Grailog: Mapping Generalized Graphs to Computational Logic

Symposium on Natural/Unconventional Computing and its Philosophical Significance, AISB/IACAP World Congress - Alan Turing 2012, 2-6 July 2012, Birmingham, UK

The Grailog User Interface for Knowledge Bases of Ontologies & Rules

OMG Technical Meeting, Ontology PSIG, Cambridge, MA, 21 June 2012

Grailog: Knowledge Representation with Extended Graphs for Extended Logics

SAP Enterprise Semantics Forum, 24 April 2012

Grailog: Towards a Knowledge Visualization Standard BMIR Research Colloquium, Stanford, CA, 4 April 2012 PARC Research Talk, Palo Alto, CA, 29 March 2012

RuleML/Grailog: The Rule Metalogic Visualized with Generalized Graphs

PhiloWeb 2011, Thessaloniki, Greece, 5 October 2011

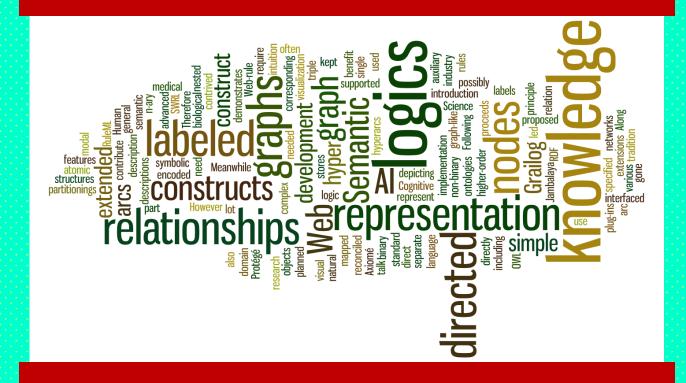
Grailog: Graph inscribed logic

Course about Logical Foundations of Cognitive Science, TU Vienna, Austria, 20 October -10 December 2008

Visualization of Data

- Useful in many areas, needed for big data
- Gain knowledge insights from data analytics, ideally with the entire pipeline visualized
- Statistical visualization → Logical visualization

Sample data visualization (http://wordle.net): Word cloud for frequency of words from BMIR abstract of this talk



Visualization of Data & Knowledge: Graphs Remove Entry Barrier to Logic

- From 1-dimensional symbol-logic knowledge specification to 2-dimensional graph-logic visualization in a systematic <u>2D syntax</u>
 - Supports human in the loop across knowledge elicitation, specification, validation, and reasoning
- Combinable with graph transformation, ('associative') indexing & parallel processing for efficient implementation of specifications
- Move towards model-theoretic semantics

 Unique names, as graph nodes, mapped directly/ injectively to elements of semantic interpretation

Grailog

Graph inscribed logic provides intuition for logic

Advanced cognitively motivated systematic

graph standard for visual-logic data & knowledge:

Features orthogonal → easy to learn, e.g. for (Business) Analytics

Generalized-graph framework as one uniform 2D syntax for major (<u>Semantic Web</u>) logics: Pick subset for each targeted knowledge base, map to/fro RuleML sublanguage, and exchange & validate it, posing queries again in Grailog

Generalized Graphs to Represent and Map Logic Languages According to Grailog 1.0 Systematics

- We have used generalized graphs for representing various logic languages, where basically:
 - Graph nodes (vertices) represent individuals, classes, etc.
 - Graph arcs (edges) represent relationships
- Next slides: What are the principles of this representation and what graph generalizations are required?
- Later slides: How are these graphs mapped (invertibly) to logic, thus specifying Grailog as a 'GUI' for knowledge?
- Final slides: What is the systematics of Grailog features?

Grailog Principles

- Graphs should make it easier for humans to read and write logic constructs via 2D state-of-the-art representation with shorthand & normal forms, from Controlled English to logic
- Graphs should be natural extensions (e.g. n-ary) of Directed Labeled Graphs (DLGs), often used to represent simple semantic nets, i.e. of atomic ground formulas in function-free dyadic predicate logic (cf. binary <u>Datalog</u> ground facts, <u>RDF</u> triples, the <u>Open Graph</u>, and the <u>Knowledge Graph</u>)
- Graphs should allow stepwise refinements for all logic constructs: <u>Description Logic</u> constructors, <u>F-logic</u> frames, general <u>PSOA RuleML</u> terms, etc.
- Extensions to boxes & links should be orthogonal

Grailog Generalizations

- Directed hypergraphs: For n-ary relationships, directed relation-labeled (binary) arcs will be generalized to directed relation-labeled (n-ary) hyperarcs, e.g. representing relational-database tuples
- Recursive (hierarchical) graphs: For nested terms and formulas, modal logics, and modularization, 'flat' graphs will be generalized to allow other graphs as complex nodes to any level of 'depth'
- Labelnode graphs: For allowing higher-order logics describing both instances and relations (predicates), arc labels will also become usable as nodes

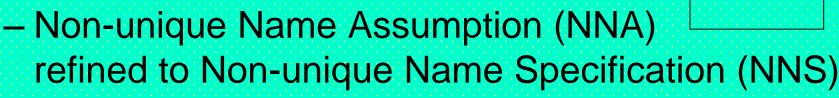
Graphical Elements: Names

Written into boxes (nodes):
 Unique (canonical, distinct) names



non-unique

- Unique Name Assumption (UNA) refined to Unique Name Specification (UNS)
- Written onto boxes (node labels):
 Non-unique (alternate, 'aka') names



• Grailog combines UNS and NNS: xNS, with x = U or N

Instances: Individual Constants with Unique Name Specifications

mapping

General: Graph (node)

unique

Examples: Graph

unique

Logic

Logic

Warren Buffett

Warren Buffett

General Electric

General Electric

US\$ 3 000 000 000

US\$ 3 000 000 000

16

Instances: Individual Constants ¹⁷ with Non-unique Name Specifications

mapping

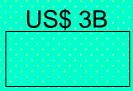
General: Graph (node)

non-unique

Examples: Graph

WB

GE



Logic (vertical bar for non-uniqueness)

Inon-unique

Logic

MB

|GE

/US\$ 3B

Graphical Elements: Hatching Patterns

No hatching (boxes): Constant
Hatching (elementary boxes): Variable

Parameters: Individual Variables

General: Graph (*hatched* node) Logic (*italics* font, <u>POSL</u> uses "?" prefix)



Examples: Graph



variable

Logic

X

Y

Α

Predicates: Binary Relations (1)

General: Graph (labeled arc)

 $inst_1 \xrightarrow{binrel} inst_2$

binrel(inst₁, inst₂)

Example: Graph

Logic

Logic

Warren Buffett Trust General Electric

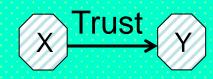
Trust(Warren Buffett, General Electric

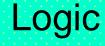
Predicates: Binary Relations (2)

General: Graph (labeled arc) Logic

binrel(var₁, var₂)

Example: Graph





Trust(X, Y)

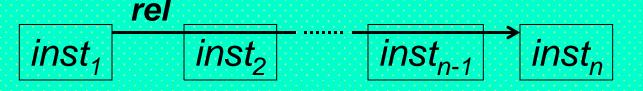
Graphical Elements: Arrows (1)

 Labeled arrows (directed links) for arcs and hyperarcs (where hyperarcs 'cut through' nodes intermediate between first and last)

Predicates: n-ary Relations (n>1)

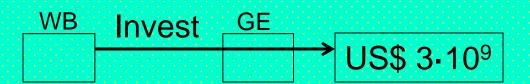
General: Graph (hyperarc)

Logic



rel(inst₁, inst₂, ..., inst_{n-1}, inst_n)

Example: Graph (n=3)



Logic

Invest(/WB, /GE, US\$ 3-10⁹)

Implicit Conjunction of Formula Graphs: Co-Occurrence on Graph Top-Level

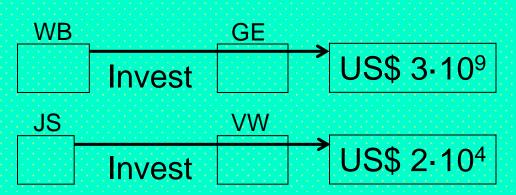
General: Graph (*m* hyperarcs)

$$inst_{1,1} \quad rel_{1} \quad inst_{1,2} \quad \cdots \quad \rightarrow inst_{1,n^{1}}$$

$$inst_{m,1} \quad rel_{m} \quad inst_{m,2} \quad \cdots \quad \rightarrow inst_{m,n^{m}}$$

Logic $rel_1(inst_{1,1}, inst_{1,2}, \dots, inst_{1,n}) \land$ $rel_m(inst_{m,1}, inst_{m,2}, \dots, inst_{m,n})$

Example: Graph (2 hyperarcs)



Logic

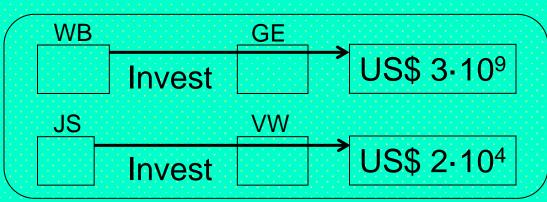
Invest(/WB, /GE, US\$ 3-10⁹) ^ Invest(/JS, /VW, US\$ 2-10⁴)

Explicit Conjunction of Formula Graphs³⁴ Co-Occurrence in (parallel-processing) And Node General: Graph (*solid+linear*) Logic

$$inst_{1,1} \text{ rel}_{1} inst_{1,2} \text{ inst}_{1,n^{1}}$$

$$inst_{m,1} \text{ rel}_{m} inst_{m,2} \text{ inst}_{m,n^{m}}$$

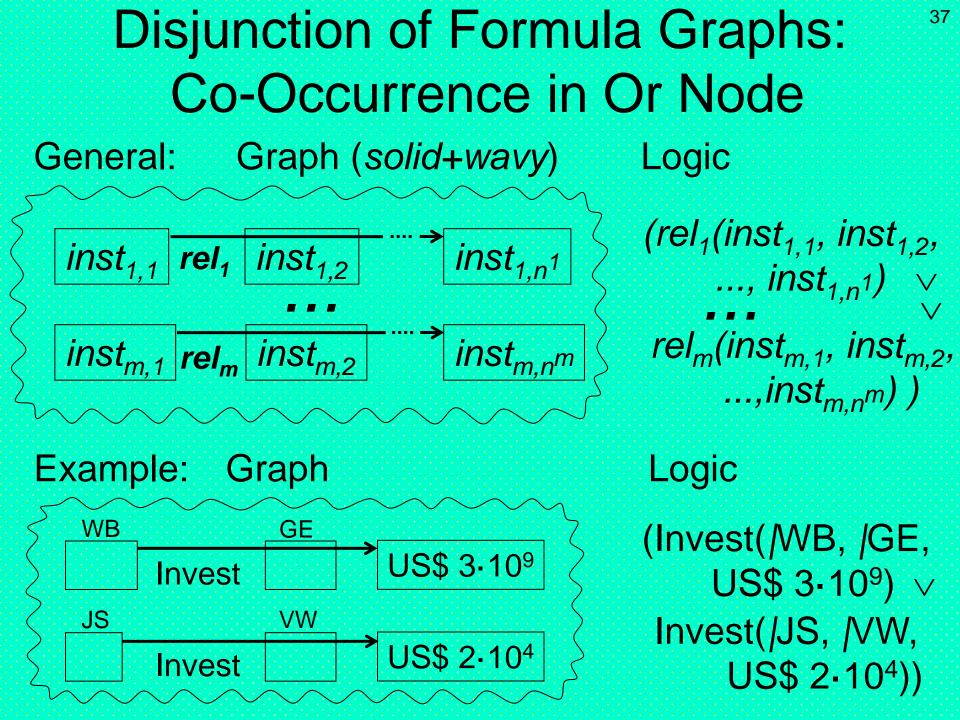
Example: Graph



 $(rel_1(inst_{1,1}, inst_{1,2}, \dots, inst_{1,n}) \land$ $rel_m(inst_{m,1}, inst_{m,2}, \dots, inst_{m,n}))$

Logic

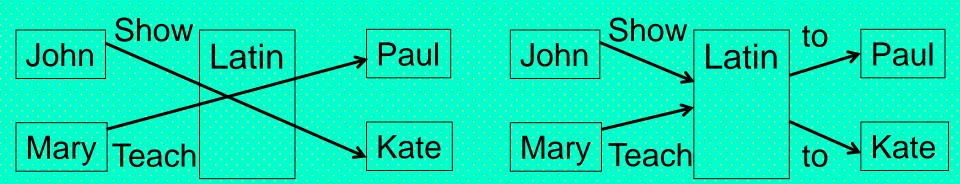
(Invest(/WB, /GE, US\$ 3⋅10⁹) ∧ Invest(/JS, /VW, US\$ 2⋅10⁴))



From Hyperarc Crossings to Node Copies as a Normalization Sequence (1)

Hypergraph (2 hyperarcs, crossing inside a node)

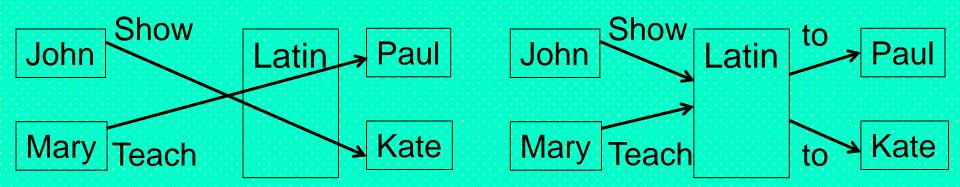
DLG (4 arcs, do <u>not</u> specify to whom Latin is shown or taught)



Symbolic Controlled English *"John shows Latin to Kate. Mary teaches Latin to Paul."* From Hyperarc Crossings to Node Copies as a Normalization Sequence (1*)

Hypergraph (2 hyperarcs, crossing outside nodes)

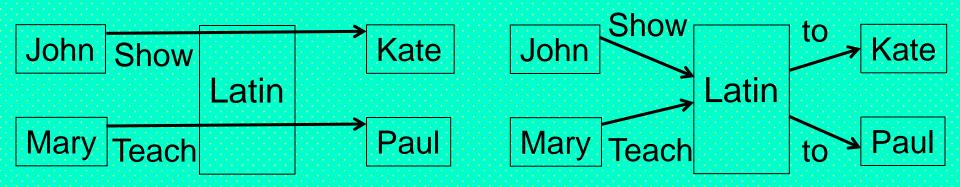
DLG (4 arcs, do <u>not</u> specify to whom Latin is shown or taught)



From Hyperarc Crossings to Node Copies as a Normalization Sequence (1**)

Hypergraph (2 hyperarcs, parallel-cutting a node)

DLG (4 arcs, do <u>not</u> specify to whom Latin is shown or taught)

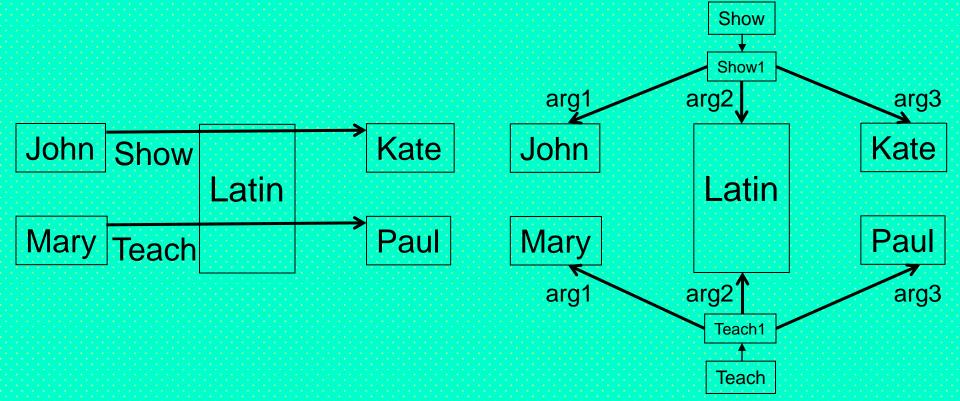


The hyperarc for, e.g., ternary Show(John,Latin,Kate) can be seen as the path composition of 2 arcs for binary Show(John,Latin) and binary to(Latin,Kate)

From Hyperarc Crossings to Node Copies — Insert on Correct Binary Reduction

Hypergraph (2 hyperarcs, parallel-cutting a node)

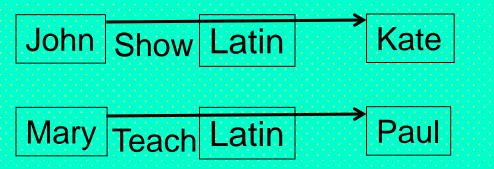
DLG (8 arcs with 4 'reified' relation/ship nodes to point to arguments)



From Hyperarc Crossings to Node Copies as a Normalization Sequence (1***)

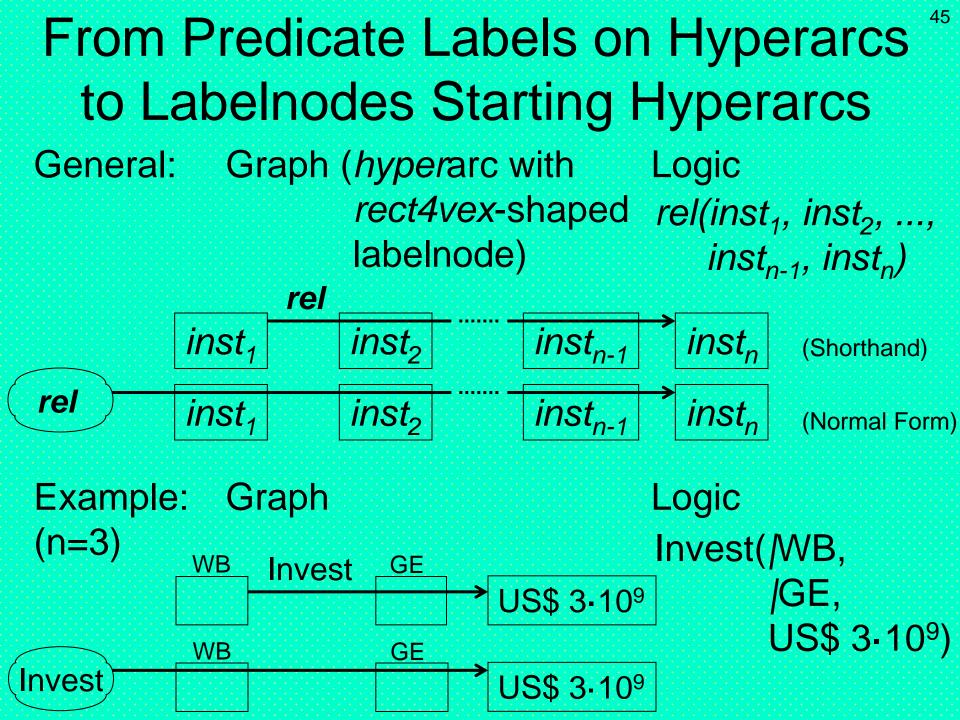
Hypergraph (2 hyperarcs, employing a node copy)

Logic (2 relations, employing a symbol copy)



Show(John, Latin, Kate) ^ Teach(Mary, Latin, Paul)

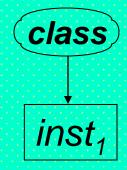
Both 'Latin' occurrences remain one node even when copied for easier layout: Having a unique name, 'Latin' copies can be merged again



Predicates: Unary Relations (Classes, Concepts, Types)

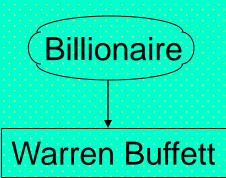
General: Graph (class applied to instance node)

HasInstance



Example: G

Graph



class(inst1)

Logic

Logic

Billionaire(Warren Buffett)

Graphical Elements: Arrows (2)

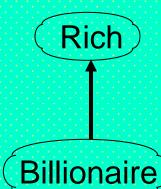
- Arrows for special arcs and hyperarcs
 - HasInstance: Connects class, as labelnode, with instance (hyperarc of length 1)
 - As in <u>DRLHs</u> and shown earlier, labelnodes can also be used (instead of labels) for hyperarcs of length > 1
 - SubClassOf: Connects subclass, unlabeled, with superclass (arc, i.e. of length 2)
 - Implies: Hyperarc from premise(s) to conclusion
 - Object-IDentified slots and shelves: Bulleted
 - arcs and hyperarcs

Class Hierarchies (Taxonomies): Subclass Relation

General: Graph (two nodes)

SubClassOf

Example: Graph



class,

class₁

(Description) Logic

 $class_1 \equiv class_2$

(Description) Logic

Intensional-Class Constructions (Ontologies): Class Intersection

General: Graph (*solid+linear* node, as for conjunction)

class₂)

Benefactor

(Description) Logic class₁ Π class₂ Π ... Π class_n

Example: Graph

Billionaire

class₁

Environmentalist

class_n

(Description) Logic Billionaire П Benefactor П Environmentalist

Intensional-Class Applications: Class Intersection

class_n

Environmentalist

General: Graph (*complex* class applied to instance node)

inst₁

Warren Buffett

class₂

Graph

Benefactor

class₁

Example:

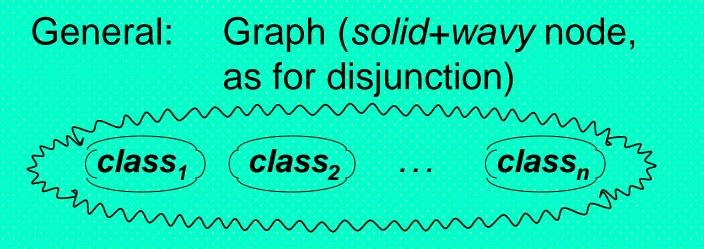
Billionaire

(xNS-Description) Logic (class₁ Π class₂ Π class_n) (inst₁)

65

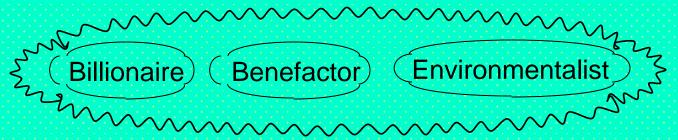
(xNS-Description) Logic (Billionaire П Benefactor П Environmentalist) (Warren Buffett)

Intensional-Class Constructions (Ontologies): Class Union



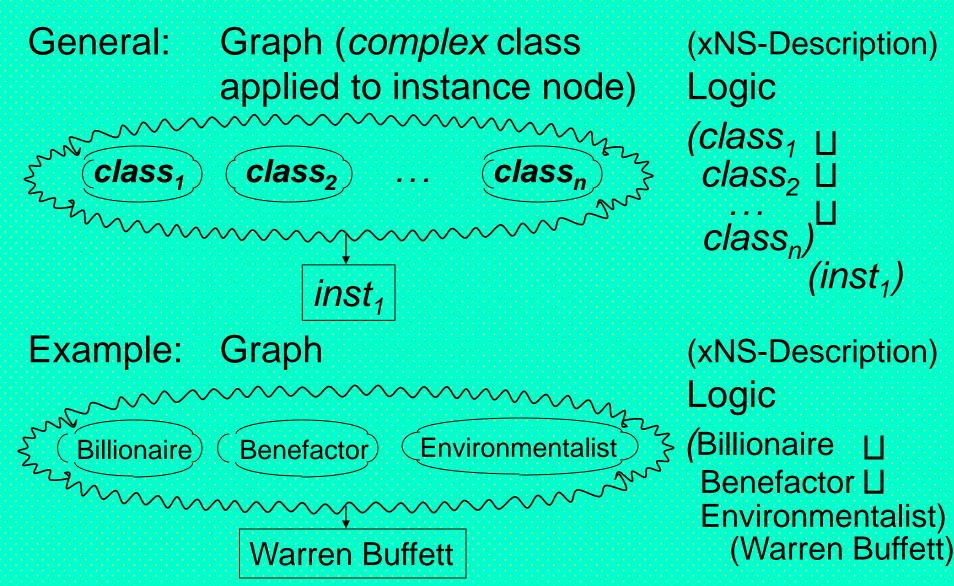
(Description) Logic class₁ ⊔ class₂ ⊔ … ⊔ class_n

Example: Graph



(Description) Logic Billionaire ⊔ Benefactor ⊔ Environmentalist

Intensional-Class Applications: Class Union



Class Hierarchies (Taxonomy DAGs): Top and Bottom

General: Top (special node)



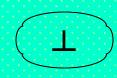
(Description) Logic

1

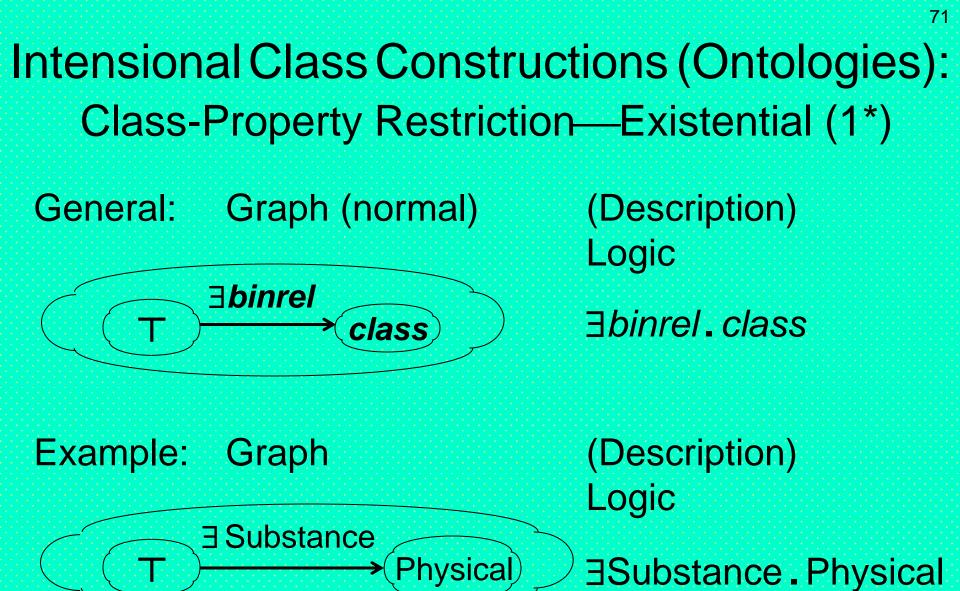
(owl:Thing)

General: Bottom (*special* node)

(Description) Logic



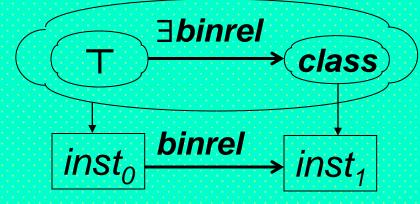
(owl:Nothing)



A kind of schema, where Top class is specialized to have (multi-valued) attribute/property, Substance, with <u>at least one</u> value typed by class Physical

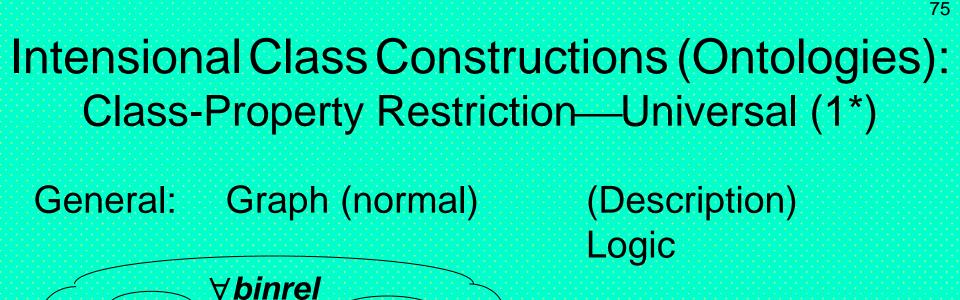
Instance Assertions (Populated Ontologies): Using Restriction for ABox—Existential (1*)

General: Graph (normal)



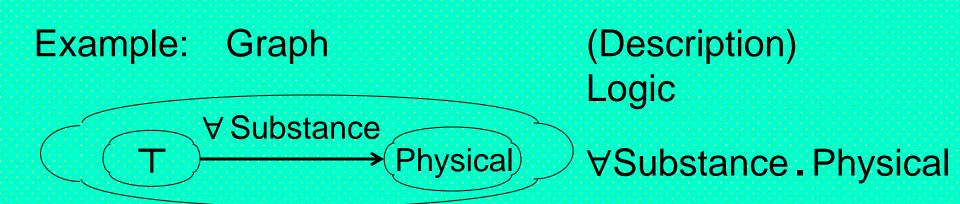
Example: Graph T Substance Physical Socrates Substance, P1 (xNS-Description) Logic $\exists binrel.class(inst_0) \land$ class(inst_1) \land binrel(inst_0, inst_1)

(xNS-Description) Logic 3Substance.Physical (Socrates) ^ Physical(P1) ^ Substance(Socrates,P1)

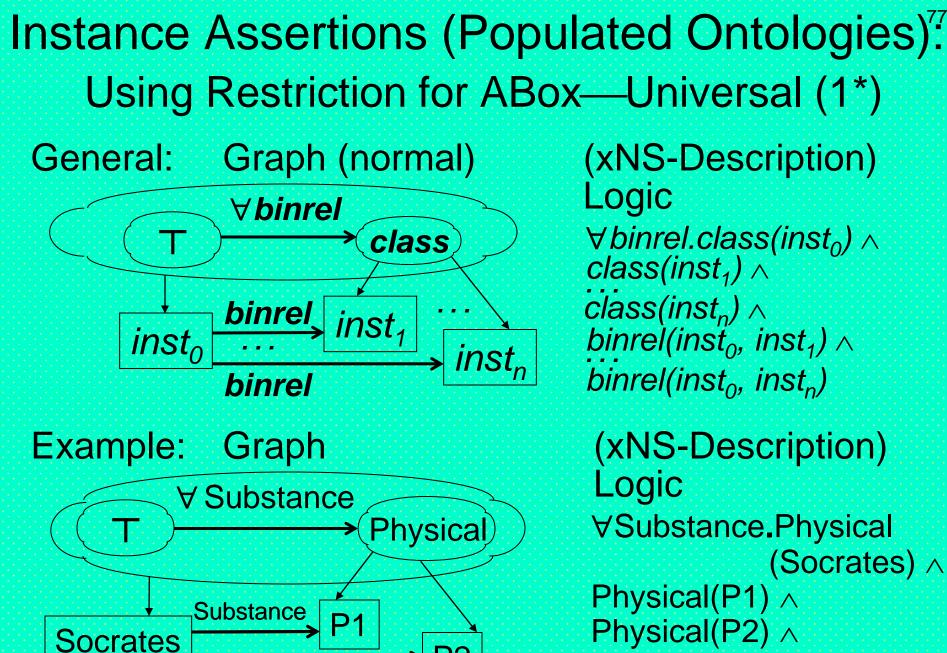


class

∀binrel.class



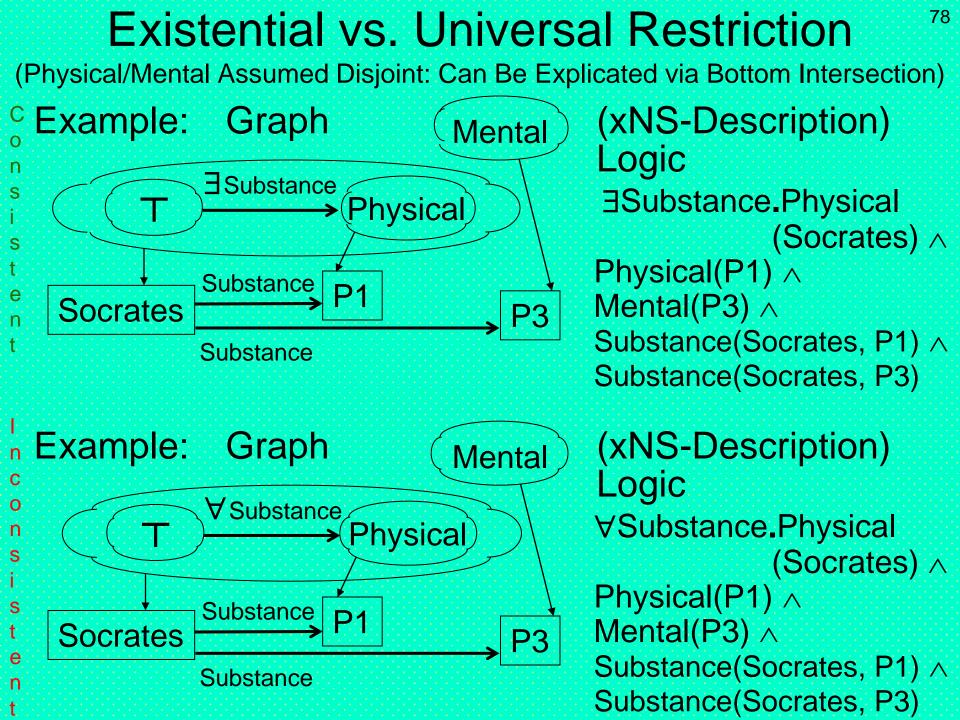
A kind of schema, where Top class is specialized to have (multi-valued) attribute/property, Substance, with <u>each</u> value typed by class Physical



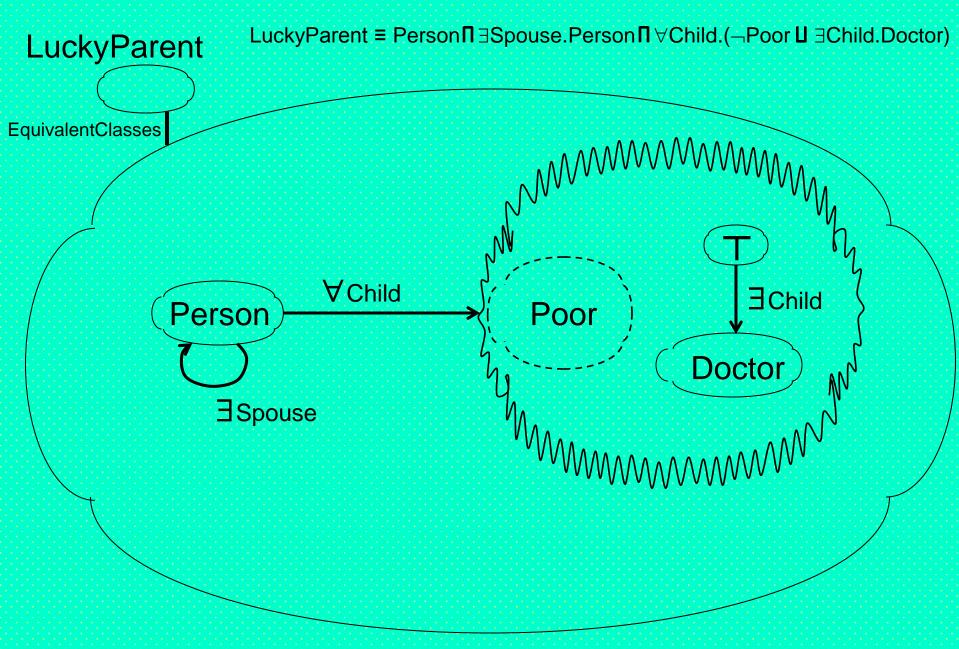
Substance(Socrates, P1) ∧

Substance(Socrates, P2)

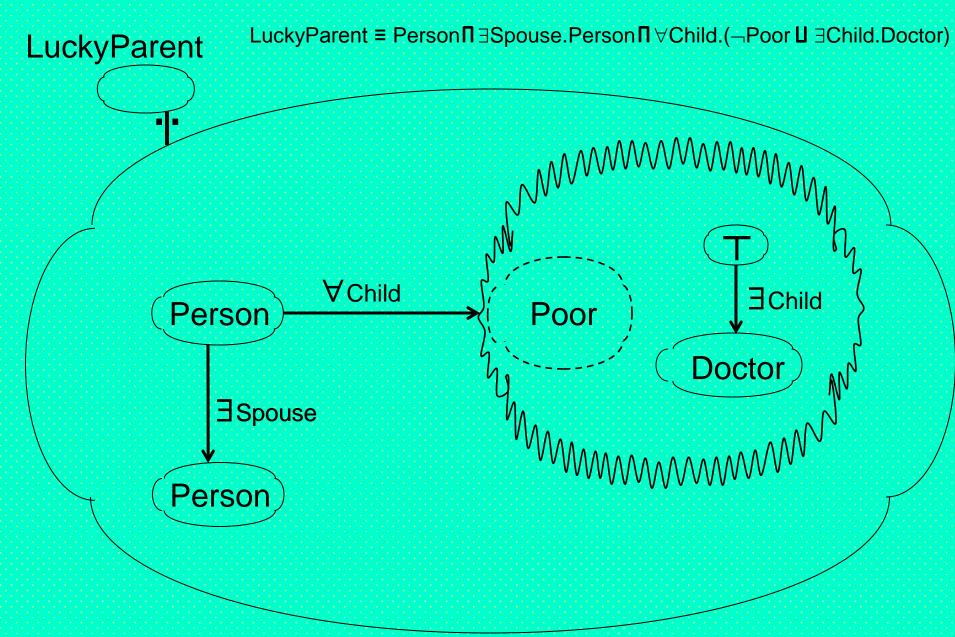
es Substance



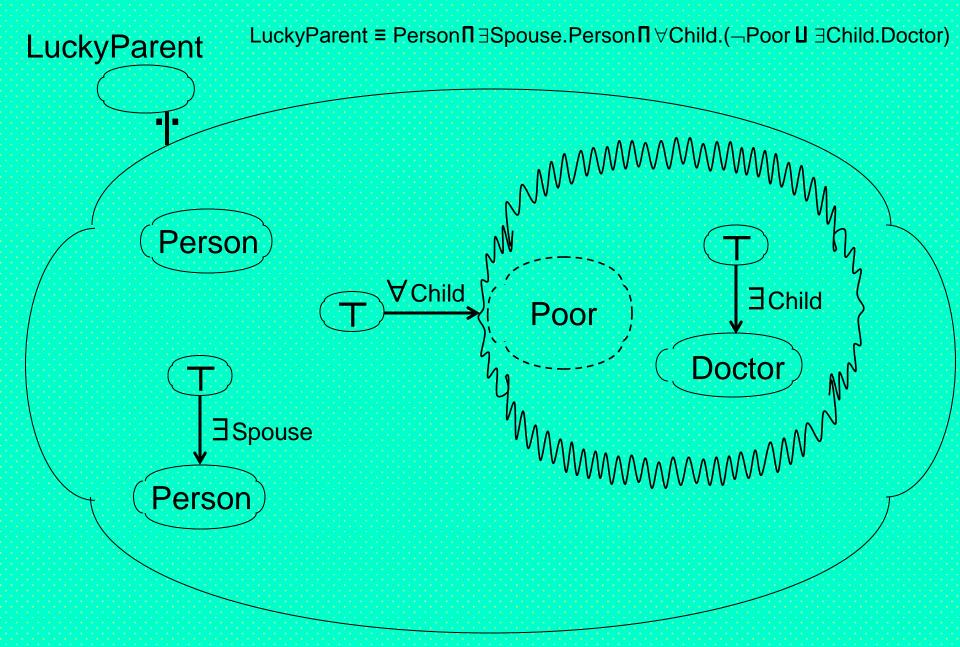
LuckyParent Example (1)



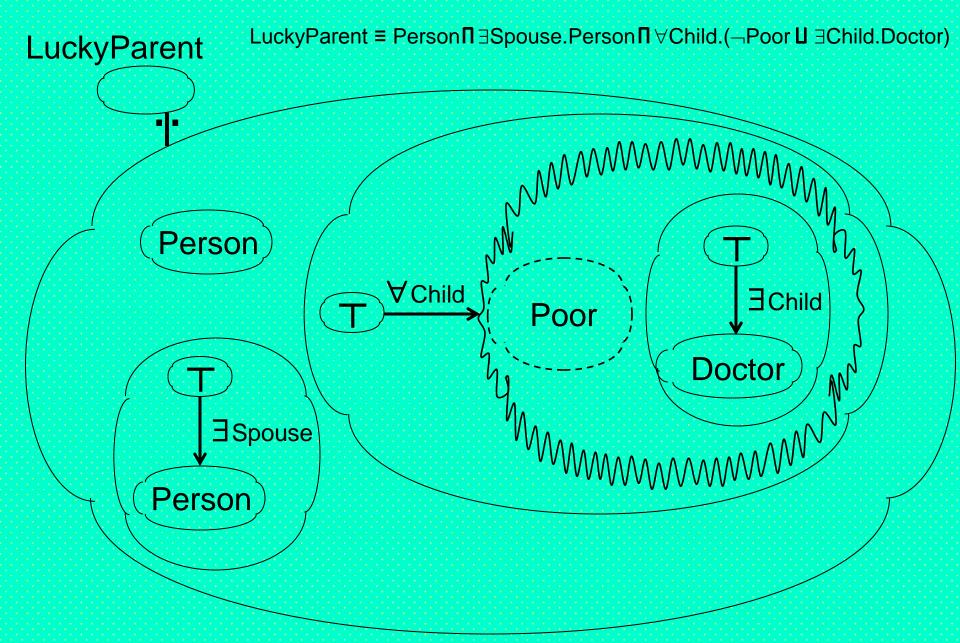
LuckyParent Example (1*)

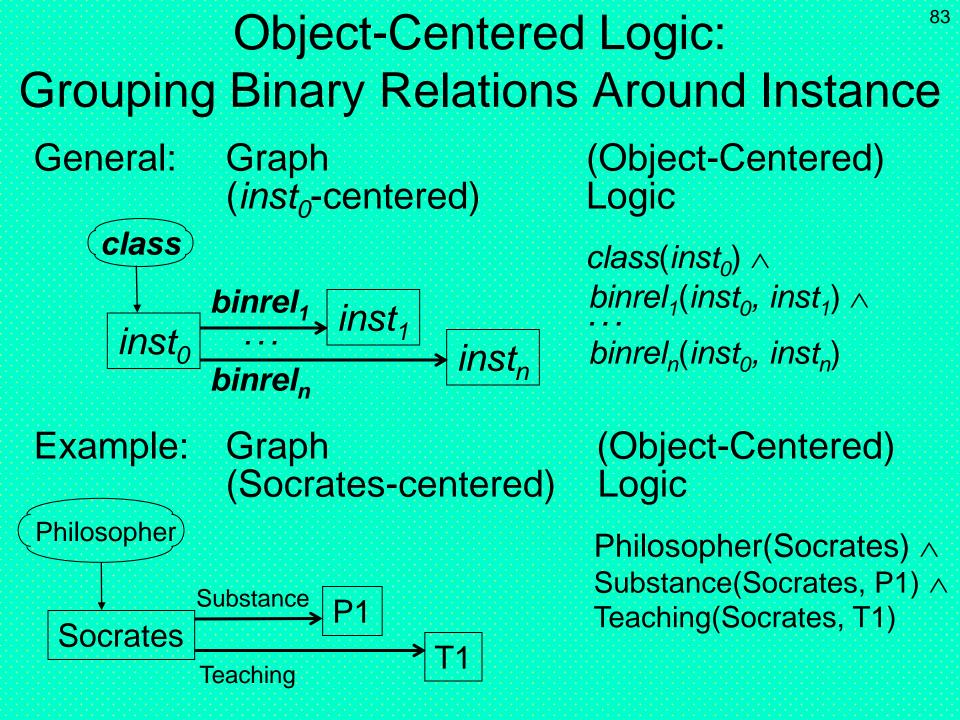


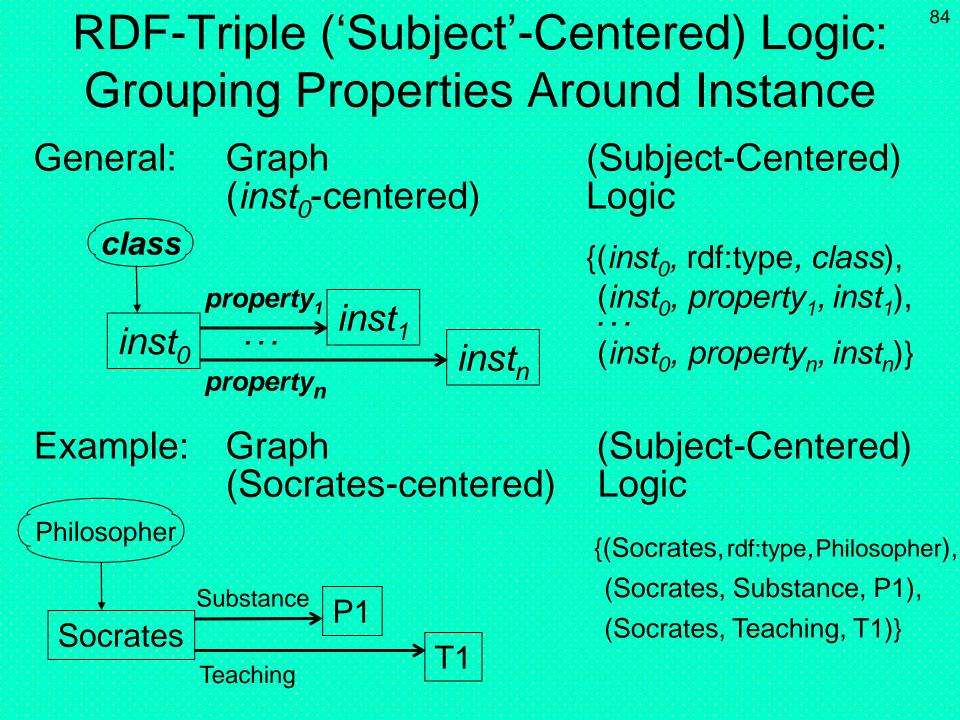
LuckyParent Example (1**)

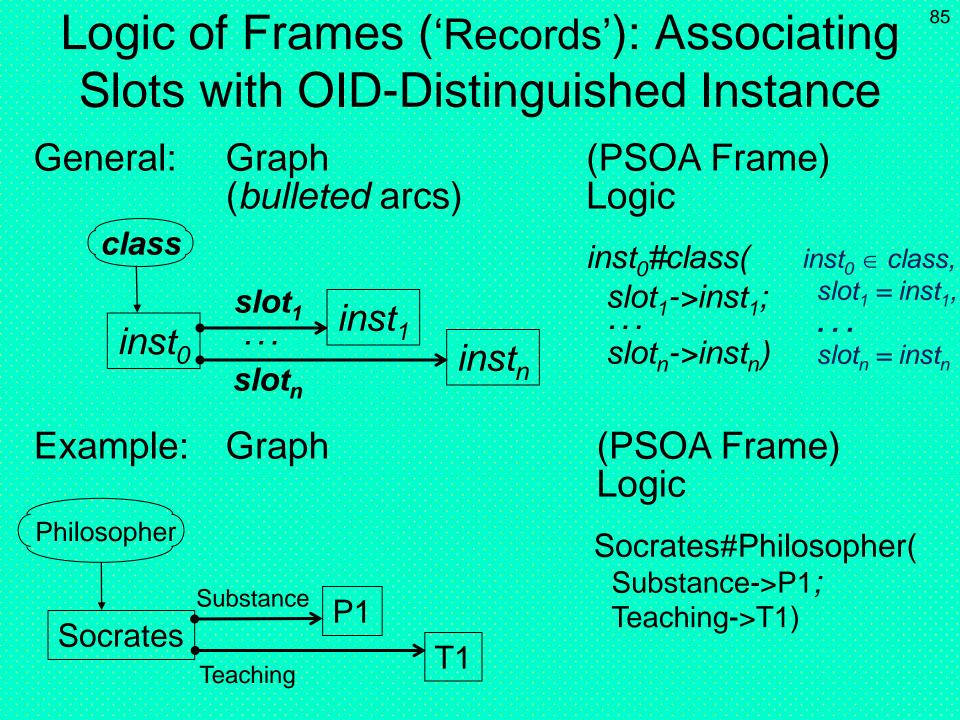


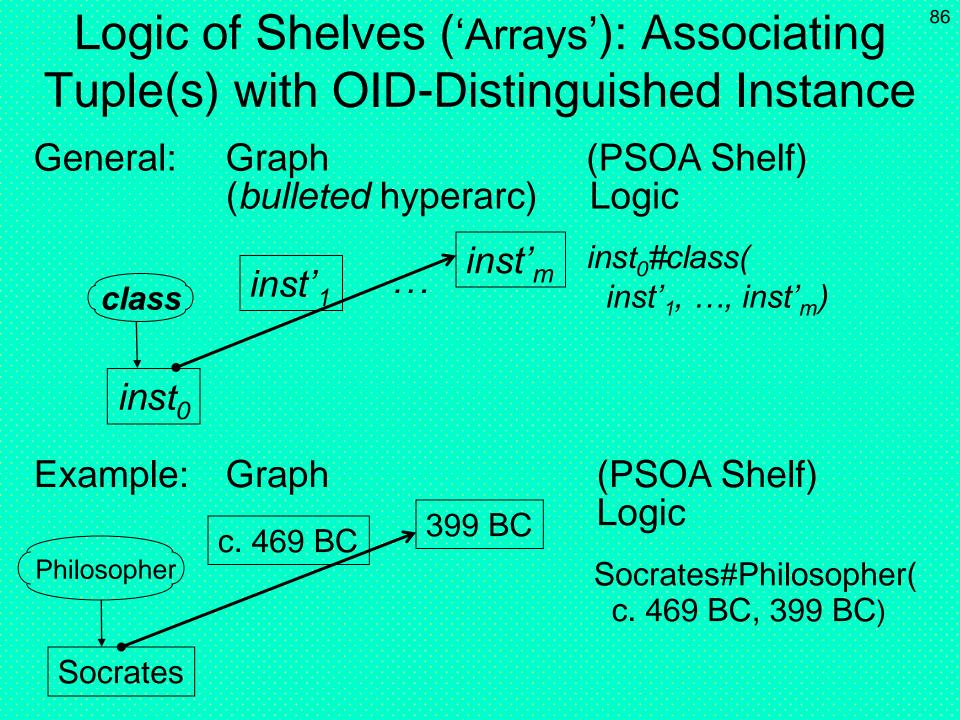
LuckyParent Example (1**)



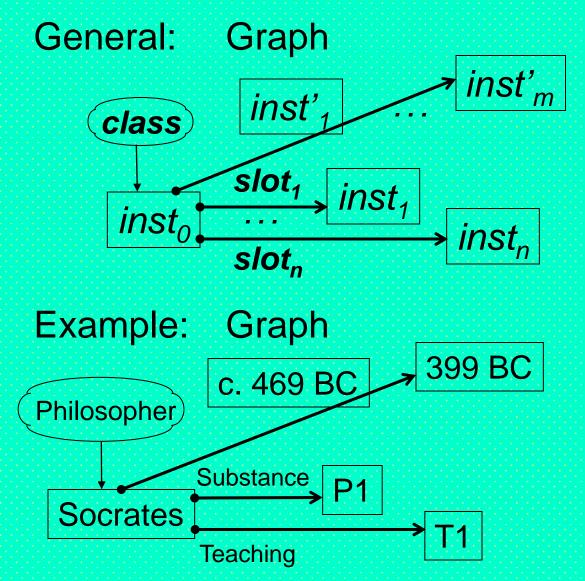








Positional-Slotted-Term Logic: Associating⁸⁷ Tuple(s)+Slots with OID-Disting'ed Instance



(PSOA Positional-Slotted-Term) Logic

inst₀#class(
 inst'₁, ..., inst'_m;
 slot₁->inst₁;
 ...
 slot_n->inst_n)

(PSOA Positional-Slotted-Term) Logic

Socrates#Philosopher(c. 469 BC, 399 BC; Substance->P1; Teaching->T1)

Rules: Relations Imply Relations (1)

General: Graph (ground, shorthand)

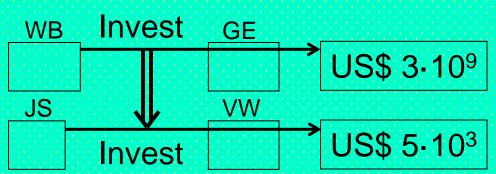
$$[inst_{1,1}] [inst_{1,2}] [inst_{1,n}]$$

$$[inst_{2,1}] [rel_2] [inst_{2,2}] [inst_{2,n}]$$

 $rel_{1}(inst_{1,1}, inst_{1,2}, \\ \dots, inst_{1,n^{1}}) \Rightarrow$

 $rel_{2}(inst_{2,1}, inst_{2,2}, ..., inst_{2,n^{2}})$

Example: Graph

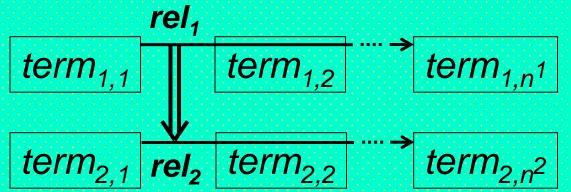


Logic

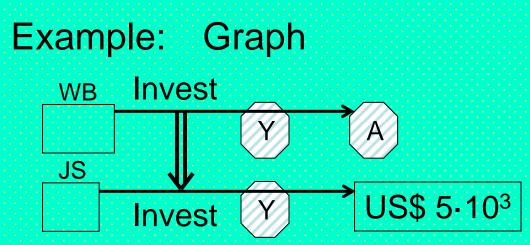
Logic

Invest(/WB, /GE, US\$ $3 \cdot 10^9$) \Rightarrow Invest(/JS, /VW, US\$ $5 \cdot 10^3$)

Rules: Relations Imply Relations (3) General: Graph (inst/var terms) Logic

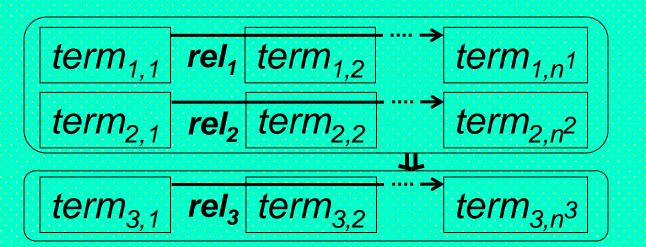


Logic $(\forall var_{i,j})$ $rel_1(term_{1,1}, term_{1,2}, \dots, term_{1,n^1}) \implies$ $rel_2(term_{2,1}, term_{2,2}, \dots, term_{2,n^2})$

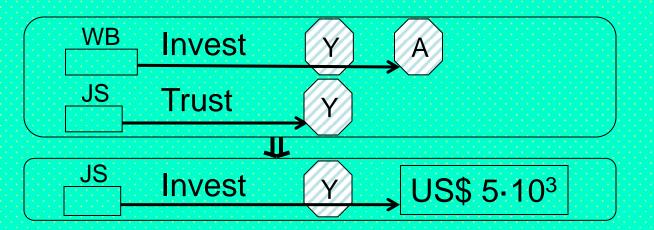


Logic $(\forall Y, A)$ Invest(/WB, Y, A) \Rightarrow Invest(/JS, Y, US\$ 5-10³)

Rules: Conjuncts Imply Relations (1*)⁹⁴ General: Graph (prenormal) Logic



Example: Graph



 $(\forall var_{i,j})$ $rel_1(term_{1,1}, term_{1,2}, \dots, term_{1,n^1}) \land$ $rel_2(term_{2,1}, term_{2,2}, \dots, term_{2,n^2}) \implies$ $rel_3(term_{3,1}, term_{3,2}, \dots, term_{3,n^3})$

Logic

 $(\forall Y, A)$ Invest(/WB, Y, A) \land Trust(/JS, Y) \Rightarrow Invest(/JS, Y, US\$ 5-10³)

Rules: Conjuncts Imply Relations (2)

Example: RuleML/XML

< Implies closure="universal"> <And> <Atom> <Rel>Invest</Rel> <Ind unique="no">WB</Ind> <Var>Y</Var> <Var>A</Var> </Atom> <Atom> <Rel>Trust</Rel> <Ind unique="no">JS</Ind> <Var>Y</Var> </Atom> </And><Atom> <Rel>Invest</Rel> <Ind unique="no">JS</Ind> <Var>Y</Var> </Atom> </Implies>

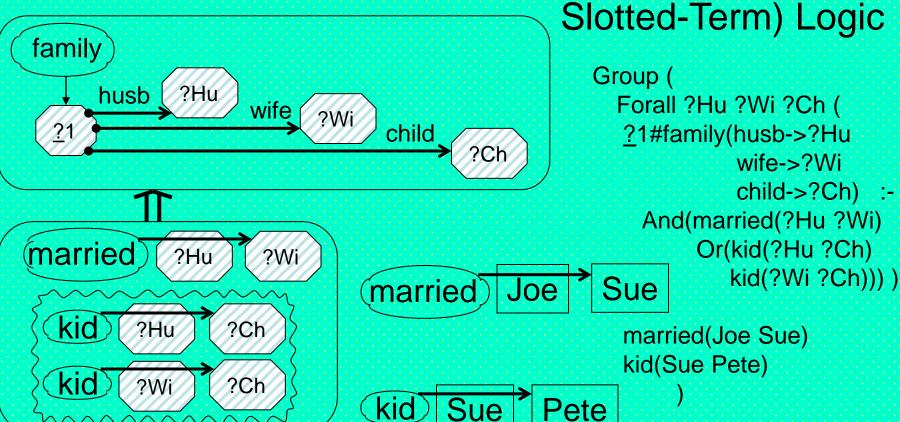
Logic $(\forall Y, A)$ $(Invest(/WB, Y, A) \land$ $Trust(/JS, Y) \Rightarrow$ $Invest(/JS, Y, US$ 5.10^3))$

Proposing an attribute unique with value "no" for NNS, and "yes" for UNS as the default

<Data>US\$ 5.10³</Data> <!- superscript "3" to be parsed as Unicode U+00B3 -->

Positional-Slotted-Term Logic: Rule-defined Anonymous Family Frame (Visualized from IJCAI-2011 Presentation)

Example: Graph

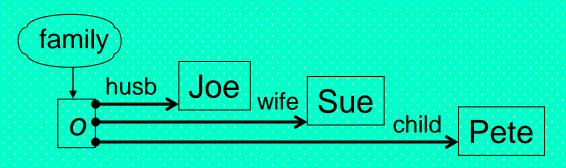


101

(PSOA Positional-

Positional-Slotted-Term Logic: Ground Facts, incl. Deduced Frame, Model Family Semantics

Example: Graph



(PSOA Positional-Slotted-Term) Logic

Group (

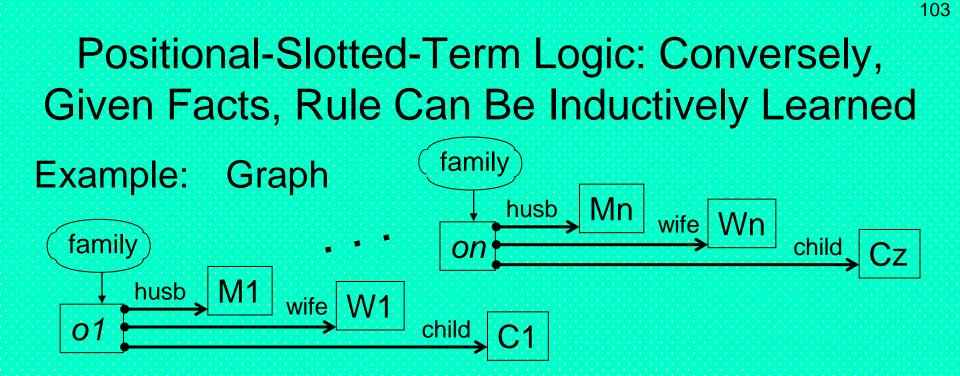
o#family(husb->Joe wife->Sue child->Pete)

Previous slide's existential variable <u>?</u>1 in rule head becomes new OID constant o in frame fact, deduced from relational facts

married(Joe Sue) kid(Sue Pete)

kid Sue Pete

For reference implementation of PSOA querying see PSOATransRun



Abstracting OID constants o1, ..., on to regain existential variable <u>?</u>1 of previous rule, now induced from matching relational and frame facts

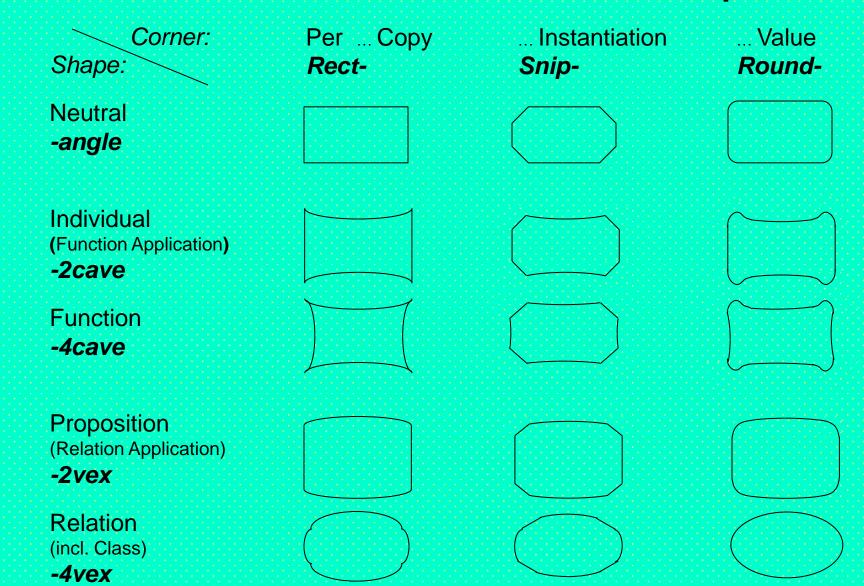


Orthogonal Graphical Features — Axes of Grailog Systematics

Box axes:

- Corners: pointed vs. snipped vs. rounded
 - To quote/copy vs. reify/instantiate vs. evaluate contents (cf. Lisp, Prolog, Relfun, Hilog, RIF, and IKL)
- Shapes (rectangle-derived): composed from sides that are straight vs. concave vs. convex
 - For neutral vs. function vs. relation contents
- Contents: elementary vs. complex nodes
- Arrow axes:
 - Shafts: single vs. double
 - Heads: triangular vs. diamond
 - Tails: plain vs. bulleted vs. colonized
- Box & Arrow (line-style) axes: solid vs. dashed, linear vs. (box only) wavy

Graphical Elements: Box Systematics¹⁰⁹ — Axes of Corners and Shapes



Graphical Elements: Boxes — Function/Relation-Neutral Shape of Angles Varied w.r.t. Corner Dimension

- Rectangle: Neutral 'per copy' nodes quote their contents

 Snipangle (octagon): Neutral 'per instantiation' nodes dereference contained variables to values from context

 Roundangle (rounded angles): Neutral 'per value' nodes evaluate their contents through instantiation of variables and activation of function/relation applications



Assuming Mult built-in function

Graphical Elements: Boxes — Concave

- Rect2cave (rectangle with 2 concave top/bottom sides): Elementary nodes for individuals (instances).
 Complex nodes for quoted instance-denoting terms (constructor-function applications)
- Snip2cave (snipped): Elementary nodes for variables.
 Complex nodes for instantiated (reified) function applications
- Round2cave (rounded): Complex nodes for evaluated built-in or equation-defined function applications
- *Rect4cave* (4 concave sides): Elementary nodes for fct's. Complex nodes for quoted functional (function-denoting) terms
 Snip4cave: Complex nodes for instantiated funct'l terms
 Round4cave: Complex nodes for evaluated functional applications (active, function-returning applications)

Graphical Elements: Boxes — Convex¹¹²

- Rect2vex (rectangle with 2 convex top/bottom sides): Elementary nodes for truth constants (true, false, unknown). Complex nodes for quoted truth-denoting propositions (embedded relation applications)
 - Snip2vex: Complex nodes for instantiated (reified) relation applications
- Round2vex: Complex nodes for evaluated relation applications (e.g. as atomic formulas) and for connective uses
 - Rect4vex: Elementary nodes for relations, e.g. unary ones (classes). Complex nodes for quoted relational (relation-denoting) terms
- *Snip4vex*: Complex nodes for instantiated relat'l terms
 Round4vex (oval): Complex nodes for evaluated relat'l applications (active, relation-returning applications)

Conclusions (1)

- Grailog 1.0 incorporates feedback on earlier versions
- Graphical elements for novel box & arrow systematics using orthogonal graphical features
 - Leaving color (except for IRIs) for other purposes, e.g. highlighting subgraphs (for retrieval and inference)
- Introducing Unique vs. Non-unique Name Specification
- Focus on mapping to a family of logics as in RuleML
- Use cases from cognition to technology to business
 - E.g. "Logical Foundations of Cognitive Science": <u>http://www.ict.tuwien.ac.at/lva/Boley_LFCS/index.html</u>
- Processing of earlier Grailog-like DRLHs studied in Lisp, FIT, and Relfun
- For Grailog, aligned with Web-rule standard RuleML: <u>http://wiki.ruleml.org/index.php/Grailog</u>

Conclusions (2)

- Symbolic-to-visual mappings implemented as Semantic Web Techniques Fall 2012 Projects:
 - Team 1 A Grailog Visualizer for Datalog RuleML via XSLT
 2.0 Translation to SVG by Sven Schmidt and Martin Koch: An Int'l Rule Challenge 2013 paper & demo introduced Grailog KS Viz
 - Team 8 Visualizing SWRL's Unary/Binary Datalog RuleML in Grailog by Bo Yan, Junyan Zhang, and Ismail Akbari: A Canadian Semantic Web Symposium 2013 paper gave an overview
- Grailog invites feature choice or combination
 - E.g. n-ary hyperarcs or n-slot frames or both
- Grailog Initiative on open standardization calls for further feedback for future 1.x versions

Future Work (1)

- Refine/extend Grailog, e.g. along with <u>API4KB</u> effort
 - Compare with other graph formalisms, e.g. Conceptual Graphs (<u>http://conceptualstructures.org</u>) and <u>CoGui</u> tool
 - Define mappings to/fro UML structure diagrams + OCL, adopting UML behavior diagrams (<u>http://www.uml.org</u>)
- Implement further tools, e.g. as use case for (Functional) RuleML (<u>http://ruleml.org/fun</u>) engines
 - More mappings between graphs, logic, and RuleML/XML: Grailog generators: Further symbolic-to-visual mappings Grailog parsers: Initial visual-to-symbolic mappings
 - Graph indexing & querying (cf. <u>http://www.hypergraphdb.org</u>)
 - Graph transformations (normal form, typing homomorphism, merge, ...)
 - Advanced graph-theoretical operations (e.g., path tracing)
 - Exploit Grailog parallelism in implementation

Future Work (2)

- Develop a Grailog structure editor, e.g. supporting:
 - Auto-specialize of neutral application boxes (angles) to function apps (2caves) or relation apps (2vexes), depending on contents
 - Auto-specialize of neutral operator boxes (angles) to functions (4caves) or relations (4vexes), depending on context
- Benefit from, and contribute to, Protégé visualization plug-ins such as <u>Jambalaya</u>/<u>OntoGraf</u> and <u>OWLViz</u> for OWL ontologies and <u>Axiomé</u> for SWRL rules
- Proceed from the 2-dimensional (planar) Grailog to a 3-dimensional (spatial) one
 - Utilize advantages of crossing-free layout, spatial shortcuts, and analogical representation of 3D worlds
 - Mitigate disadvantages of occlusion and of harder spatial orientation and navigation
- Consider the 4th (temporal) dimension of animations to visualize logical inferences, graph processing, etc.