Factors affecting ontology development in ecology

Data Integration in the Life Sciences (DILS2005)

20-22 July 2005, San Diego, USA

C. Maria Keet

KRDB Research Centre
Faculty of Computer Science
Free University of Bozen-Bolzano
Italy
keet@inf.unibz.it

Overview

- Introduction, motivation, comparative analysis
- Development considerations
- Translation STELLA to ontology elements
- Experimental results and discussion
- Conclusions

Overview

- **Introduction, motivation, comparative analysis**
- Development considerations
- Translation STELLA to ontology elements
- Experimental results and discussion
- Conclusions

Introduction: CS/IT Motivation

- Extensive use of modelling in ecology, but not much shared (depending on sub-discipline). [e.g. *Eco. Mod.* and *Env. Mod. & Sw*]
- Models used with independent software tools (DB and other applications)
- `Legacy code' (procedural), moving toward more OO, and sparingly ontologies [Argent04]
- Requirement for (re-)analysis to upgrade legacy SW [Keller&Dungan99], develop new SW to meet increasing complexities and rising demands.

  use the opportunity to create a more durable, yet computationally usable, shared, agreed upon representation of the knowledge about reality.

Introduction: Salient characteristics of ecology and comparative analysis

- Flow of 'components', tight coupling between endurants and perdurants, event-centred
- Granularity and scope of the subject domain(s)
- Modelling tradition

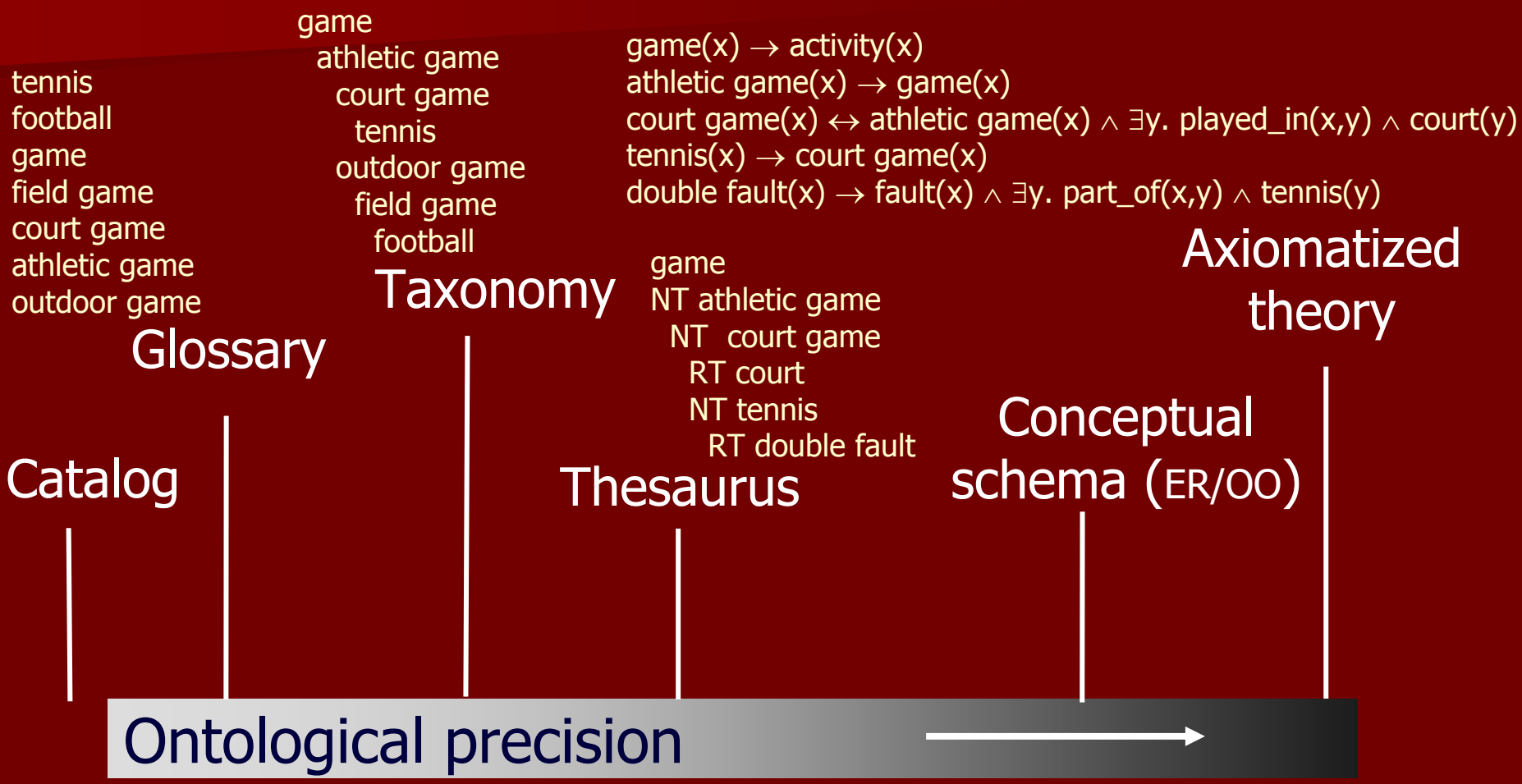
Overview

- Introduction, motivation, comparative analysis
- **Development considerations**
- Translation STELLA to ontology elements
- Experimental results and discussion
- Conclusions

Development considerations: some problems with taxonomies

- (in)formal taxonomy cannot represent the same amount of knowledge as in STELLA
  - Subject domain, context
  - Restriction of properties and relations
  - DAG-Edit lacks support for Racer
    (to check consistency, classify taxonomy)
- Multiple inheritance

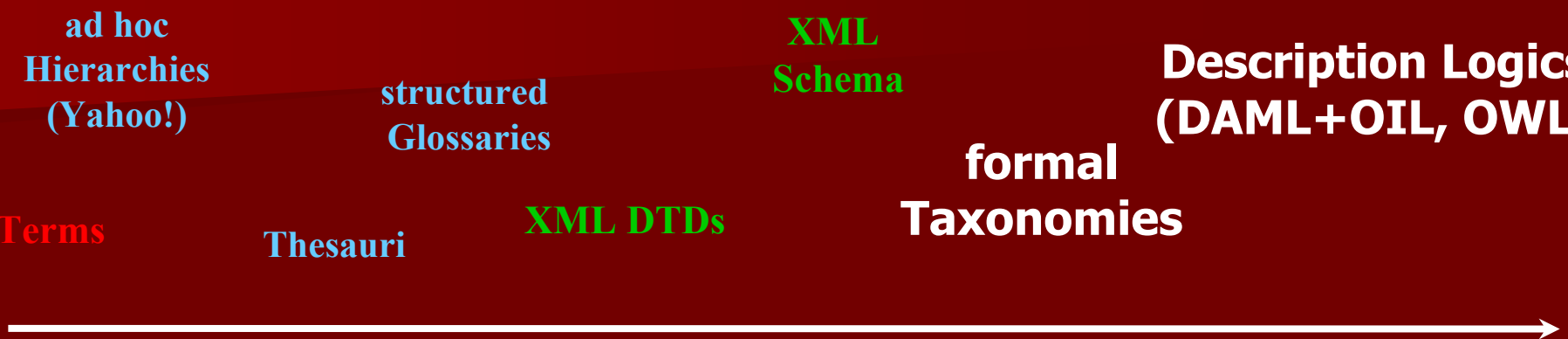
Development considerations: Levels of Ontological Precision



precision: the ability to catch all and only the intended meaning (for a logical theory, to be satisfied by intended models)

[Gangemi04]

Development considerations: languages



Glossaries & Data Dictionaries

Thesauri, Taxonomies

MetaData, XML Schemas, & Data Models

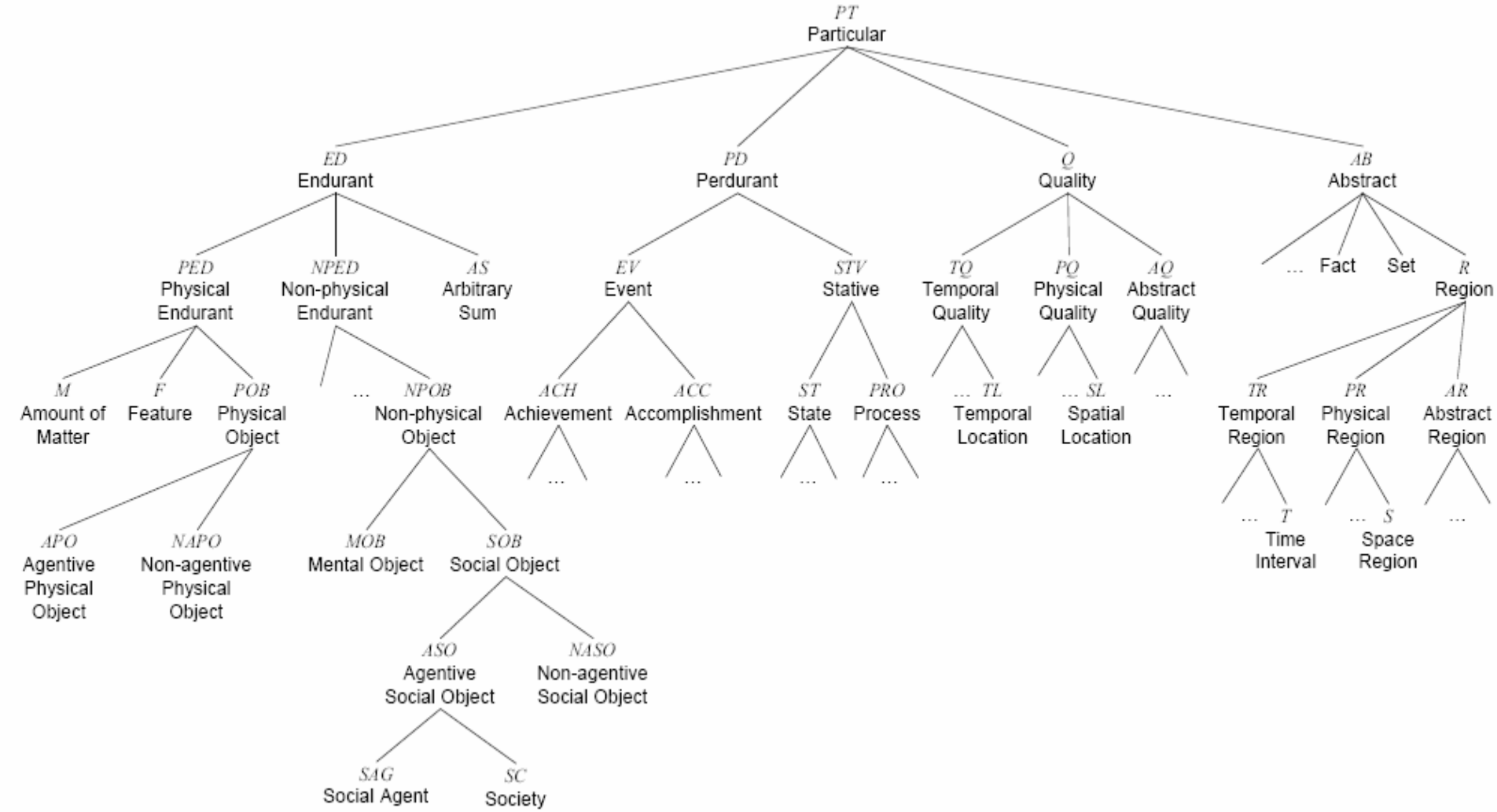
Formal Ontologies & Inference

Based on [Gruninger]

Development considerations: formal & foundational

- Formal, DL-based, Protégé: Maximum expressiveness without losing computational completeness. But:
  - Need self-restraint for property creation
  - Protégé restricts naming of the entities
  - Limited options for visualisation
- Use of foundational ontology for guidance during ontology development: *Don't reinvent the wheel if you don't want to*

Development considerations: formal & foundational



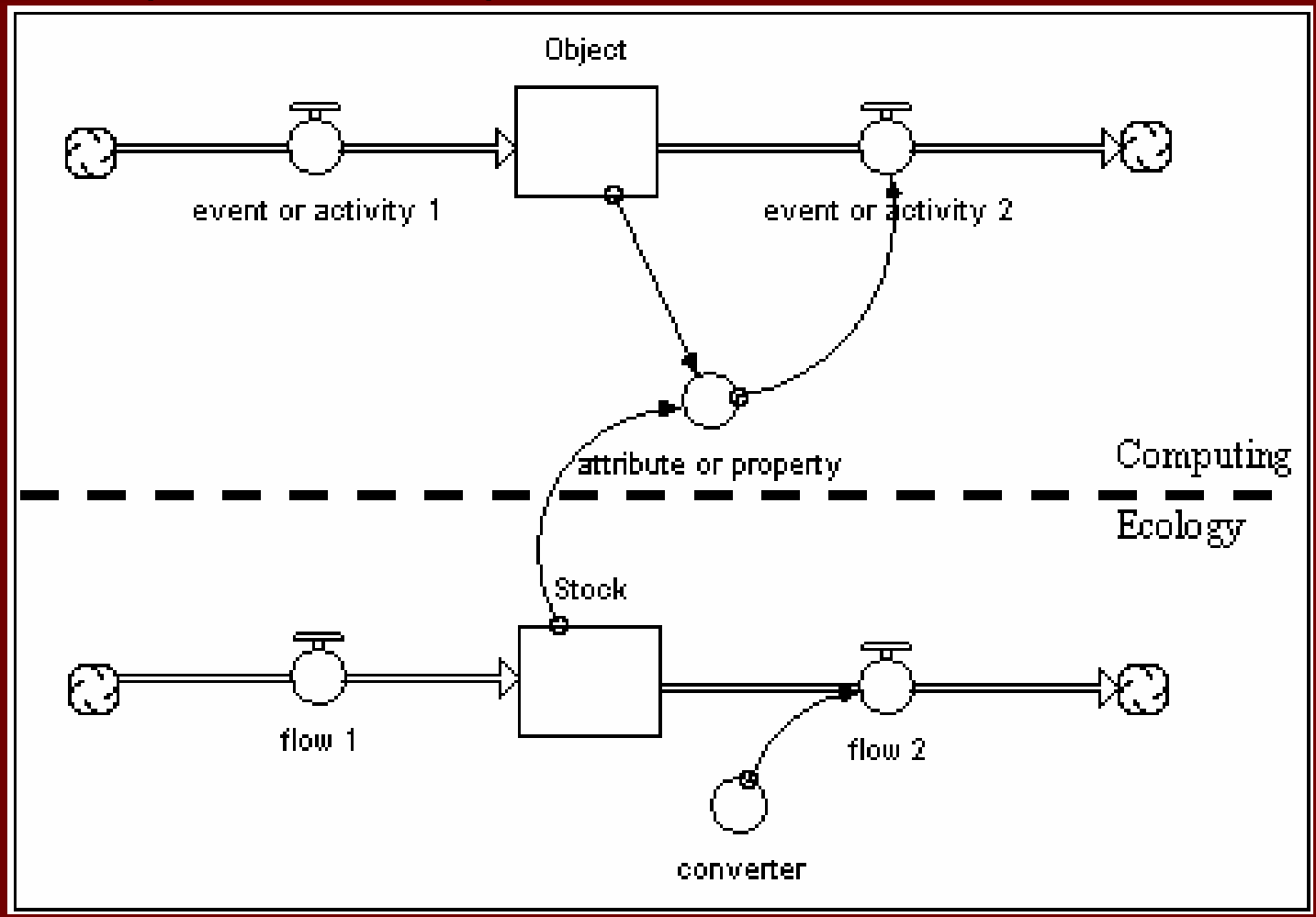
[Masolo et. al 03]

Overview

- Introduction, motivation, comparative analysis
- Development considerations
- **Translation STELLA to ontology elements**
- Experimental results and discussion
- Conclusions

Translation (1/3)

- Key aspects in the ecological model: Flow, Stock, Converter, Action Connector.



Translation (2/3)

- A **Stock** correspond to a noun (particular or universal)
- **Flow** to verb
- **Converter** to attribute related to Flow or Stock
- **Action Connector** relates the former

- ***Object*** is candidate for an endurant
- ***Event_or_activity*** for a method or perdurant
- Converter maps to ***attribute_or_property***
- Action Connector candidate for relationship between any two of Flow, Stock and Converter.

Translation (3/3)

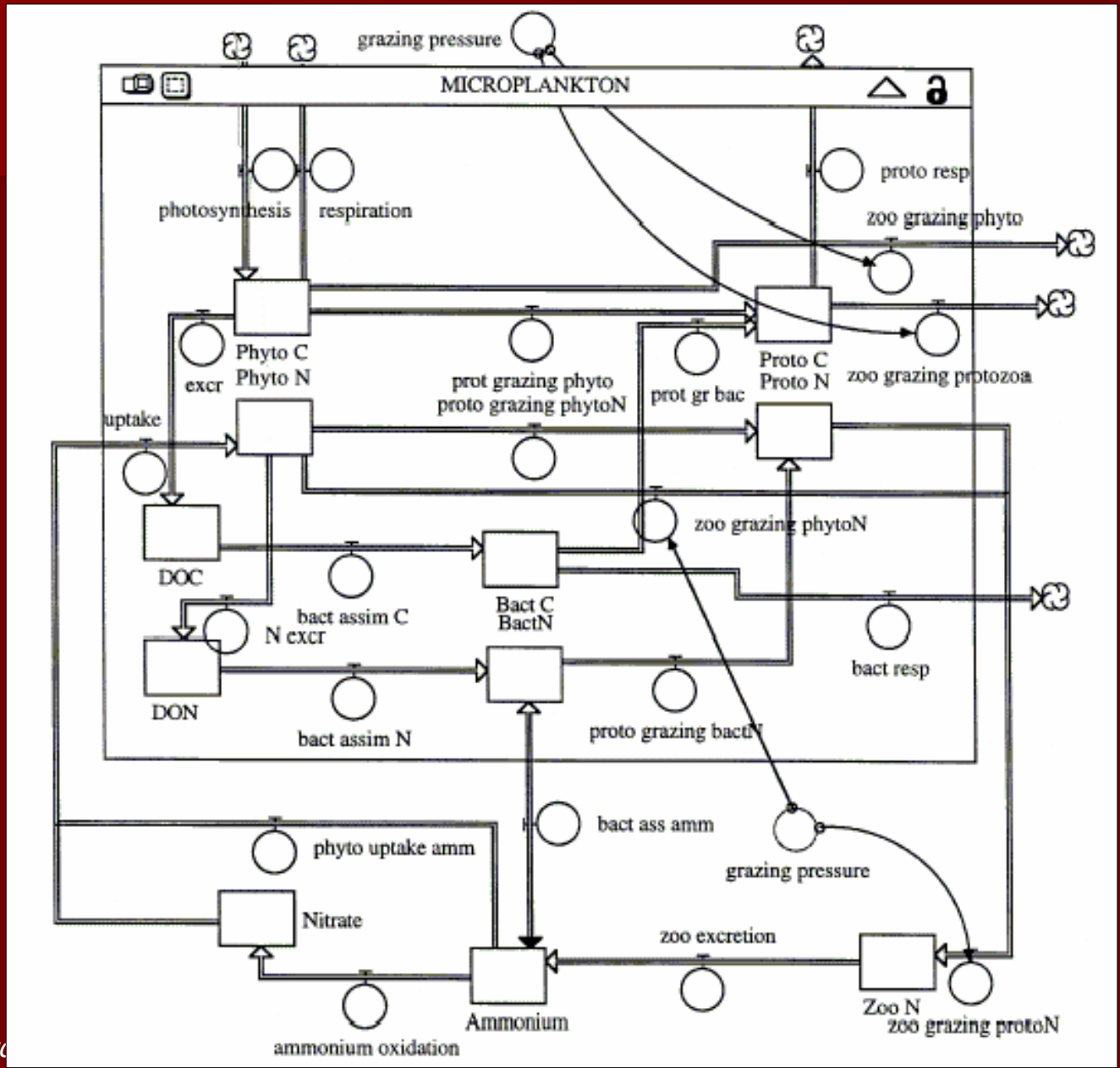
- $\forall x\,((\text{Stock}(x) \leftrightarrow \text{Entity}(x)) \rightarrow \text{ED}(x))$
- $\forall x\,((\text{Flow}(x) \leftrightarrow \text{Entity}(x)) \rightarrow \text{PD}(x))$
- $\forall x\,((\text{Converter}(x) \leftrightarrow \text{Entity}(x)) \rightarrow (\text{Q}(x) \vee \text{ST}(x)))$
- $\forall x\,(\text{ActionConnector}(x) \leftrightarrow \text{Relationship}(x))$

ED = EnDurant, PD = PerDurant, Q = Quality, ST = State

Overview

- Introduction, motivation, comparative analysis
- Translation STELLA to ontology elements
- Development considerations
- **Experimental results and discussion**
- Conclusions

Microbial Loop (1/5)

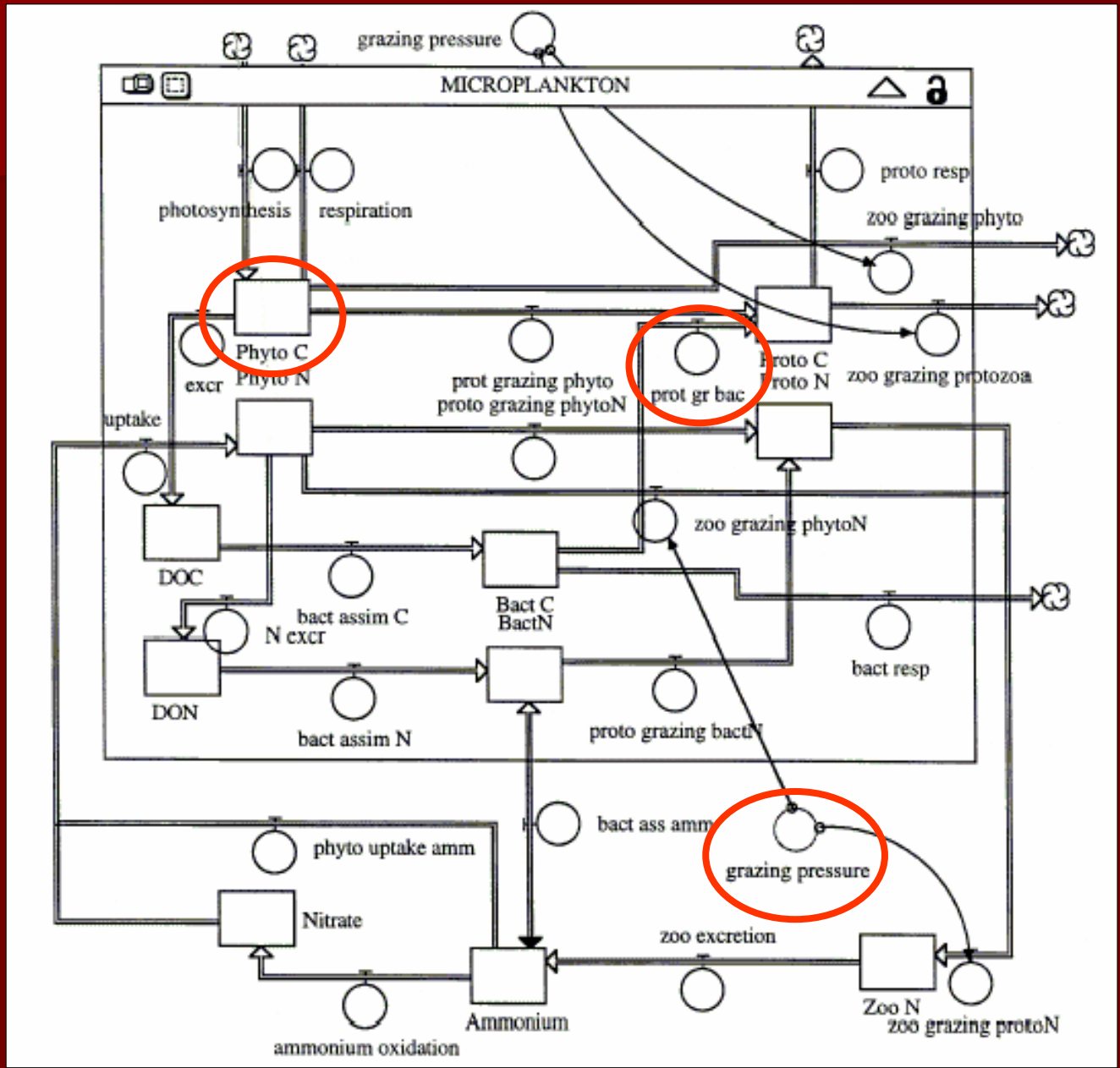


[Tett&Wilson00]

Microbial Loop (2/5)

- Aim: to test translations with a real STELLA model
- ML's initial mapping to ontological categories contain 38 STELLA elements: 11 Stock/ED, 21 Flow/PD, 2 Converters/ST, 4 Action Connectors/Relationships.
- The *MicrobialLoop* ontology has 59 entities and 10 properties.
- Increase due to including DOLCE categories and implicit knowledge of ML that is explicit in *MicrobialLoop*.

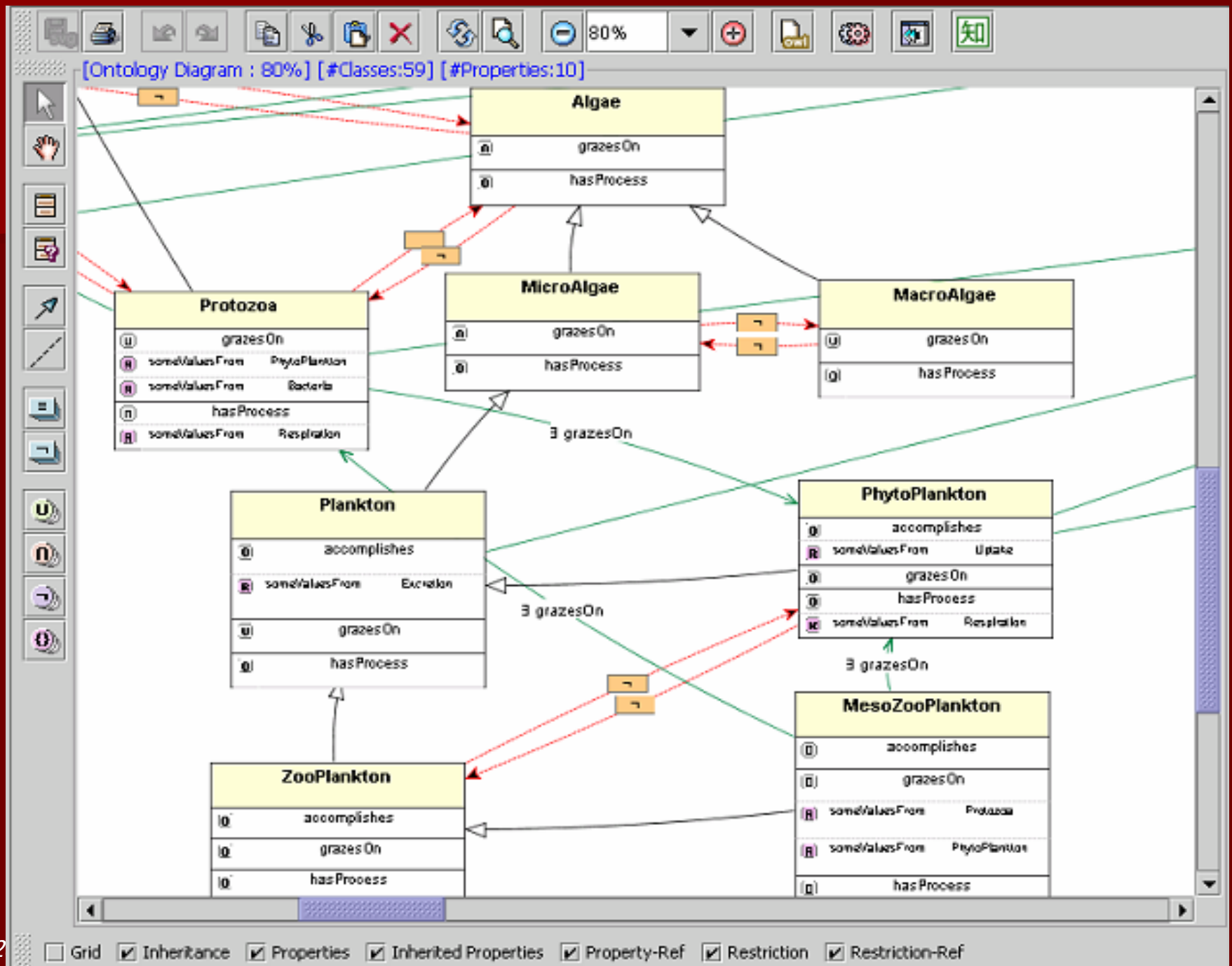
Microbial Loop (3/5)



Microbial Loop (4/5)

Stock	DOLCE category	Comments
Phyto C	**NAPO**	**Phyto C = phytoplankton organic carbon. Phytoplankton is an APO, but 'phyto C' is *part* of the APO: only the organic carbon of the phytoplankton, not the organism as an active agent as such**
Phyto N	NAPO	Phyto N = phytoplankton nitrogen
DOC	NAPO	DOC = detrital organic carbon. Detritus is an ED with no unity, thus an amount of matter (M), but here, like with the organisms, there is focus on only a *part* of the NAPO
Nitrate	NAPO	Dissolved nitrate. Molecules are non agentive physical objects.
Flow		
Photosynthesis	PRO	To phytoplankton N
Respiration	PRO	From phytoplankton N
Prot gr bac	**PRO**	**Protozoa that are grazing on the Bacterial C**
Converter		
Grazing pressure	**ST**	**Acts on a PRO affecting the process of grazing; 'grazing pressure' is there (might reach zero), hence a ST.**
Action connector		
"1"	Yes	Acts on the mesozooplankton grazing on the protozoa, and acts on the mesozooplankton grazing on the phytoplankton: relation *hasGrazingPressure*

Microbial Loop (5/5)



Discussion

- Formalising ecological natural, functional and integrative concepts:
  - aids comparison of theories,
  - makes the implicit explicit, and more expressive than other modelling practices therefore useful:
  - points to ambiguous sections,
  - part of/extra tool for doing science,
  - importance ontology maintenance, comparisons
- Modular, backbone or all-encompassing ontology/ies
- With the mappings, a quicker bottom-up development of ecological ontologies

Overview

- Introduction, motivation, comparative analysis
- Translation STELLA to ontology elements
- Development considerations
- Experimental results and discussion
- **Conclusions**

Conclusions

- Taxonomies insufficiently expressive compared to existing ecological modelling techniques
- Perspective of flow in ecological models cannot be represented adequately in a taxonomy.
- More comprehensive semantics of formal ontologies.
- Formalised mapping between STELLA and ontology elements facilitates bottom-up ontology development and has excellent potential for semi-automated ontology development.
- STELLA as intermediate representation, widely used by ecologists and is translatable to a representation usable for ontologists.

Concl. & ongoing/future research

- Ontology development for ecology is close to being part of ecological research that through the formalized representation of the knowledge more clearly points to lacunas and suggestions for further research in ecology, thereby aiding hypothesis generation.
- We are currently extending this research with ontology development and management aspects such as modularisation, granularity, and ontology integration.

Thank you!