

Ubiquitous Data

Gavin Bierman, Cambridge

Peter Buneman, Edinburgh

Philippa Gardner, Imperial

Databases

The topic of databases emerged from two key requirements:

- simple abstraction of structured data
- robust, efficient implementation of large quantities of data.

Successful solution: relational databases

Distributed databases provide an early, practical example of mobile code. The number of data locations is typically very small.

With **ubiquitous databases**, the emphasis is now on **scale**.

Examples

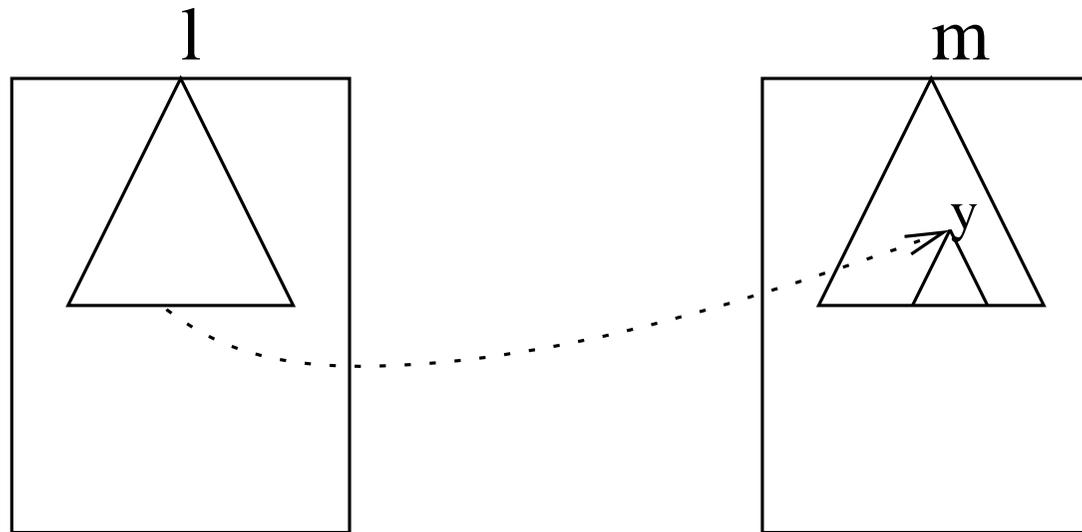
Bioinformatics databases: 500 public databases, many commercial databases. Key features:

- evolving structure of data
- few original data sources
- continual monitoring of data sources for new information.

Future health care: patients 'own' their records such as genetic sequence data. A researcher correlating the occurrence of cardiovascular condition with genetic structure must

- access potentially millions of databases
- move **trusted** code to source data.

Semi-structured Data



Manipulation of such data is not straightforward. We must study

- query languages: **XQuery**
- pattern-matching languages: **C# with XML types**
- storage and optimisation techniques
- provenance and annotation.

Processes

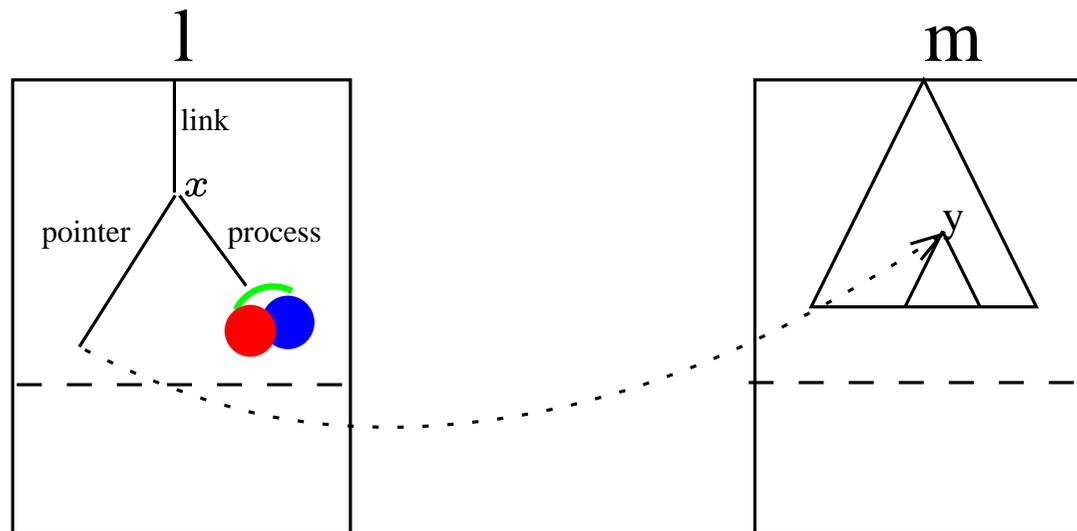
Process calculi provide primitives for describing and analysing global distributed infrastructure, focusing on

- process migration
- process interaction
- private channel communication.

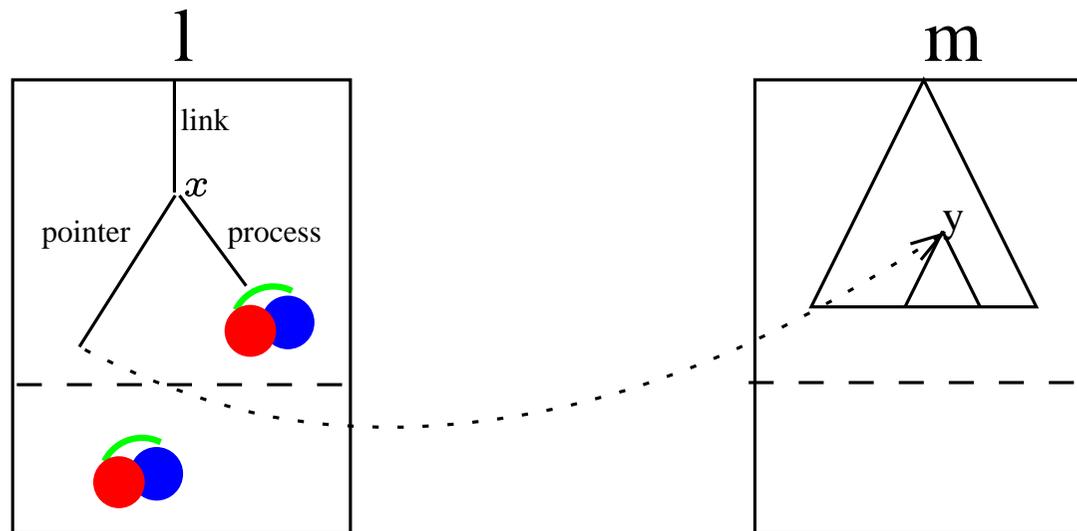
Example applications include

- languages supporting distributed programming with process mobility: polyphonic C#, BPEL
- description and analysis of authentication protocols
- typed processes to ensure fine-grained resource access control.

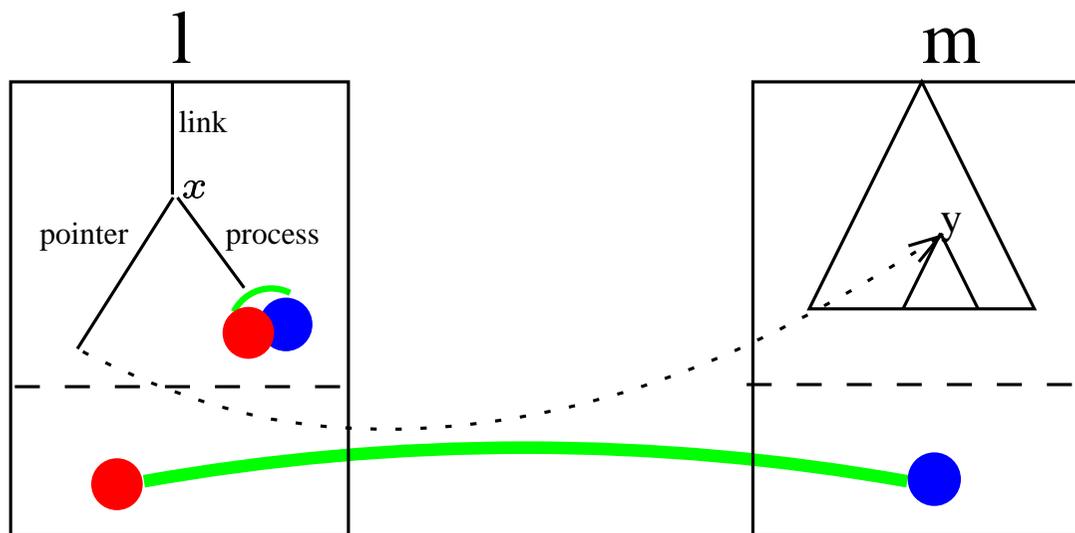
Integrating Data and Processes



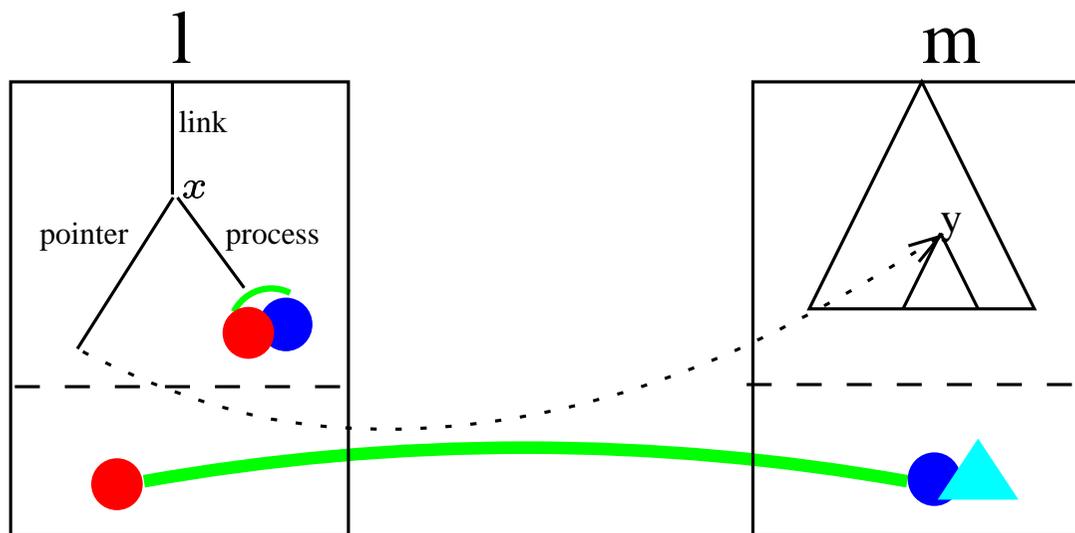
Integrating Data and Processes



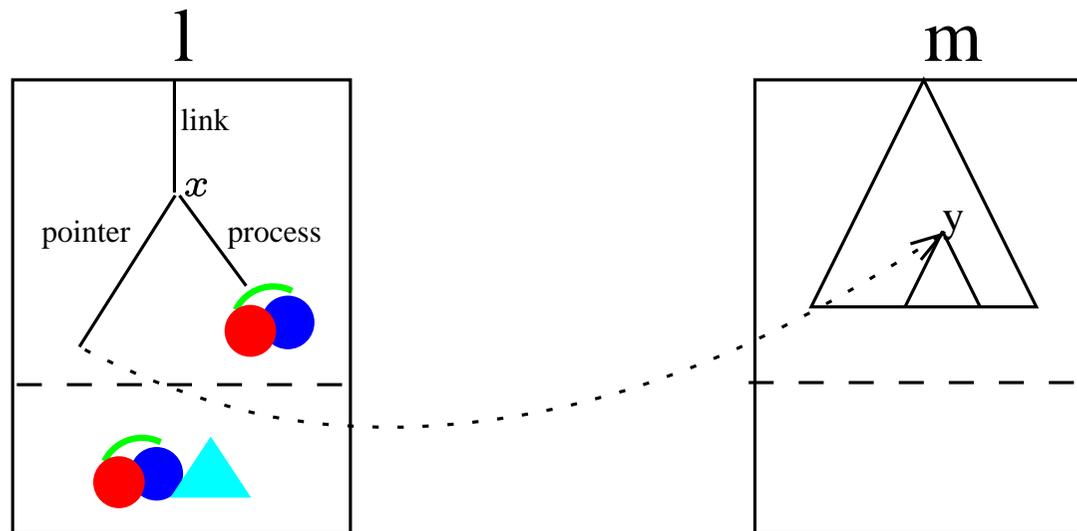
Integrating Data and Processes



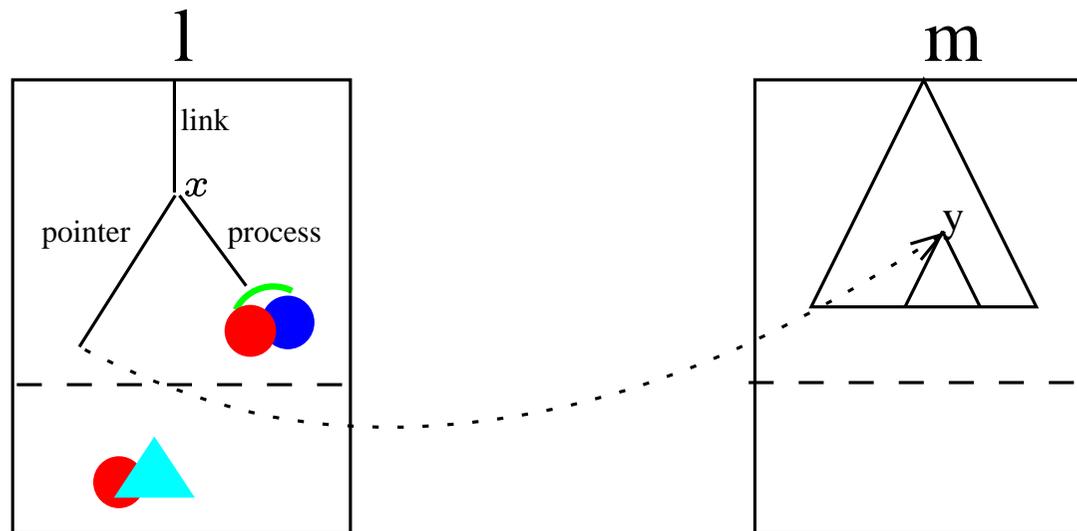
Integrating Data and Processes



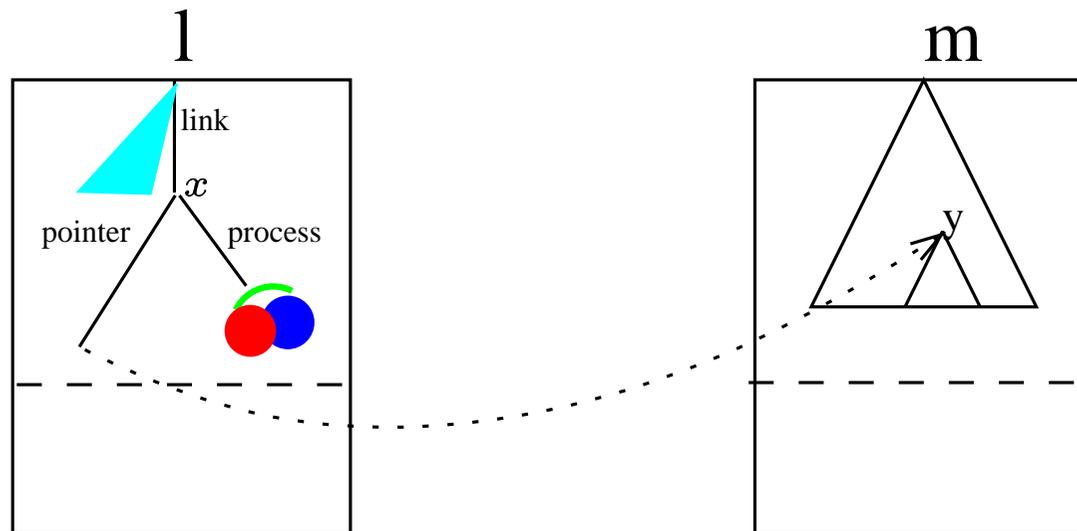
Integrating Data and Processes



Integrating Data and Processes



Integrating Data and Processes



15 Year Challenge

To find a sound theoretical underpinning for languages and data associated with the Web.