Cover Languages

Given

- a set of selected words: a.x.y.z.w.b, and
- the schema \{a.x.y.z.w.b, a.r.s.t\}.

We are checking words from the schema only against the set of selected words.

It's possible to use a.(x|y|z|w)*.b as an alternative set of selected words. Why?
Path Expressions With Iterated Wild Cards

- An abstraction mechanism over structural details.

Examples:

- XPath: `//para`
- AOP: `call(void n()) && cflow(call(void m()))`
- AP: `from A to B`
Path Expression

- from a to b
- Schema: \{a.x.y.z.w.b, a.r.s.t\}
- APE: \ a \cdot \Diamond^* \cdot b
- EAPE: \ a \cdot (x|y|z|w|r|s|t)^* \cdot b
- SEL = EAPE \cap \text{Schema} = \ a \cdot x \cdot y \cdot z \cdot w \cdot b
Question

Applicability to XML processing?
WYSIWYG Semantics

- Wild cards can be replaced only with symbols not mentioned in the path expression.
- Efficiency (Determinism).
- Non-ambiguity.
- Modularity. (Order Guarantee)

  - BR → LoB → Bus → LoP → Passenger → Pass → BR.
  - BR → Passenger.
CDAP

- `<book>` ...
- `<author>` `<name>` `<first>` `<\first>`
- `<last>` `<\last>` ...

- `<first>` & `<last>` are interchangeable.

- Print author first name followed by last name regardless of their order in the XML document.