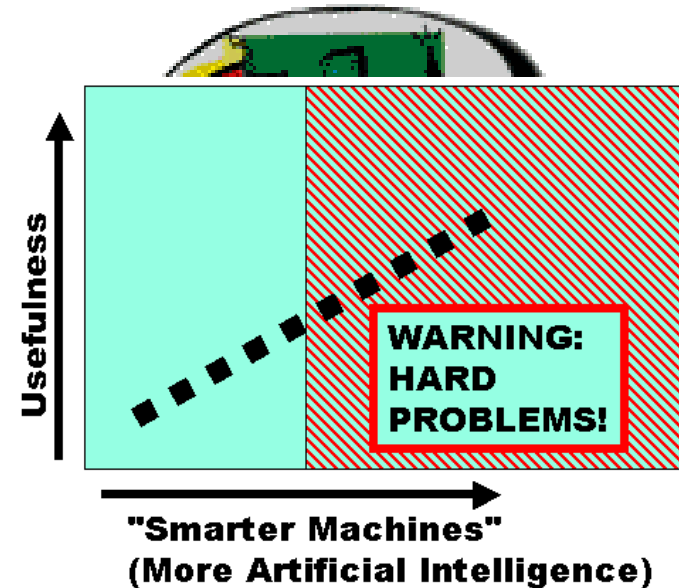


Knowledge-Driven Content Processing and the Semantic Web

Ian Horrocks

What is the Semantic Web?

- Not clear what the Semantic Web is or will be
- Hype seems to suggest that it means: “semantics + web = AI”
- I prefer to think of it as meaning: “semantics + web + AI = more useful web”



Semantic Web architecture

- Current architecture includes (at least):
 - Semantic **annotations** added to web resources
 - **Ontologies** provide vocabulary for annotations
 - Ontology derived **semantics** improve (machine) “understanding” of web content
- What does it mean to “understand” web content?
 - Having some notion of **meaning** (semantics) that goes beyond surface structure
 - The ability to **recognise similarities** based on meaning
 - The ability to **derive** additional (implicit) meaning

What does it mean for multimedia?

- **Identifying** multimedia resources based on (e.g.) content
 - E.g., Images of David Beckham
- **Recognising** relationships between apparently diverse resources
 - E.g., Images of professional sportspersons
- **Deriving** additional implicit information
 - E.g., team player/individual
- **Answering** complex queries
 - Female sportsperson whose sport is not swimming or a ball game



How will it work?

- Requires some form of reasoning
 - Ontology languages such as OWL are based on **logic**
 - Can use **automated reasoning** to
 - Make implicit information explicit
 - Answer complex queries
- Reasoning services should be
 - **Fast**
 - e.g., for query answering over very large data sets
 - **Reliable**
 - e.g., for machine understanding (no human in the loop)

Language requirements?

- Specific **features** may be important for multimedia
 - Spatial reasoning, temporal reasoning, ...
- Can **extend** languages like OWL, but
 - Extensions should be based on thorough investigations of theoretical and practical implications
 - Need to balance expressiveness against ability to provide infrastructure such as fast and reliable reasoning services
 - Try to find “low-cost” solutions, e.g., time-stamping instead of fully fledged temporal logic

No point in developing new languages without any means to support applications that use them