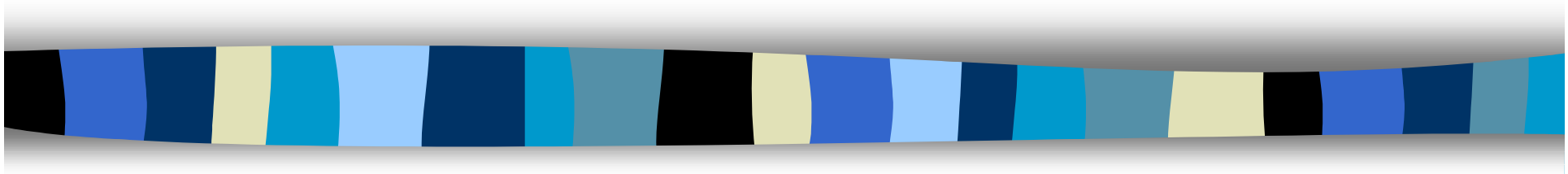
On the spanning hypothesis for EDI semantics



Steve Kimbrough
The Wharton School
University of
Pennsylvania
kimbrough@wharton.upenn.edu

Scott Moore
University of Michigan
Business School
samoore@umich.edu
www-personal.umich.edu/
~samoore/research/flbc/

The problem

- When sending a message, it is difficult to determine if the recipient is interpreting the message in the way in which the sender intended
- Applications
  - EDI (standard e-commerce)
  - Agent communication languages

Needed solution

- A semantics
  - A theory of what languages need to say (and what they already say)
- A language
  - A lexicon
  - Rules of expression formation

Not just any language

- One that spans the domain
  - It says what needs to be said
- It is parsimonious
  - Doesn't have an infinite number of predicates
- These are conflicting goals

The lexicon

- Controlled vocabulary
  - Verbs (deliver, order, bill)
  - Thematic roles
- Open vocabulary of nouns & descriptors
  - Think of ‘catalog items’

Thematic roles?

- General helper predicates used to qualify the meaning of verbs
- Examples
  - Agent, Theme, Goal, Source, Location, Instrument

Open vocabulary?

- Questions & concerns
  - How open?
  - How updated?
  - How maintained?
  - How referenced?
- These are unanswered questions
  - We're addressing them through 'proof by example'

So, our research addresses...

- Defining an appropriate language
  - Come up with appropriate rules for expression formulation
  - Settle on the appropriate set of thematic roles
  - Come up with a useful and realiztic means of maintaining the open vocabulary
- Testing it out with realistic examples

Examples we're exploring

- In this paper
  - Compaq's EDIFACT messages they use with business partners that are posted on the Web
- Otherwise
  - KQML and other agent communication languages
  - Other EDIFACT and X.12 message sets

Example 1 (simple form)

ordering(e) ∧ Agent(e, cpq) ∧
Addressee(e, vend) ∧ Theme(e, [item1]) ∧
Cul(e, '970101') ∧ original(e) ∧
referenceNumber(e, 'P1M24987E') ∧
name(cpq, 'Compaq') ∧
Type(item1, 'PRODUCT') ∧
description(item1,
'Description info goes here') ∧
quantity(item1, 100) ∧ units(item1, 'PCE') ∧ …

Example 2 (EST form)

$po(e) \wedge \underline{Speaker}(e, s) \wedge \underline{Addressee}(e, r) \wedge$
$\underline{Theme}(e, (e_1 | e_2)) \wedge \underline{Cul}(e, t)) \wedge$

$(H(e) \leftrightarrow$
$(delivering(e_1) \wedge \underline{Agent}(e_1, r) \wedge \underline{Goal}(e_1, s) \wedge$
$\underline{Theme}(e_1, g) \wedge \underline{Sake}(e_1, e) \wedge unit(e_1, g, u) \wedge$
$quantity(e_1 , g, q) \wedge \underline{Cul}(e_1 , t_1) \wedge$
$t_1 <= T_1)) \wedge ...$

[this could all be in XML (or EDIFACT…) if desired]

Findings: Concerning EDI

- EDI messages are not clearly defined
  - This is not a surprise
- Messages can follow the standard but still be idiosyncratic to trading partners
- Complexity of standard makes it very difficult to construct a valid message

Findings: Concerning research program

- We've introduced it and described it
- We're encouraged but…
- These are our first baby steps