

EDI, XML, and the Transparency Problem in Electronic Commerce

Steven O. Kimbrough

University of Pennsylvania, Philadelphia, PA, USA,
kimbrough@wharton.upenn.edu

Abstract. Standard (that is, long-standing and currently much in use) EDI protocols (including the X12 and EDIFACT series) have repeatedly been criticized for poor design, confusing or absent semantics, and much else. Most of these criticisms are indeed on the mark. The main conclusions that the critics have drawn are also correct: business-to-business e-commerce is expensive and difficult to set up and maintain, because of shortcomings in the design concepts underlying standard EDI. Something must be done, but what?

Central to the problem is the fundamental question of semantic transparency: When A sends B a message, how does B's machine know what the message is about, what it means? Given proper standards, message meanings are determined and computers can be programmed to act appropriately to the intended message meanings. The complaint against EDI has been that proper standards cannot be made because of the misguided way in which the EDI standards are designed.

Proponents of XML have been touting XML's strengths and claiming that they overcome, or can overcome, the semantic transparency problem in e-commerce. In support of this claim, proponents point to the DTDs (or similar devices) that any XML/EDI solution would use. The claim is that semantic transparency is/can be achieved through the DTDs.

In this paper I argue that indeed the DTD mechanism offers a kind of progress on the semantic transparency problem, but that it cannot provide anything approaching a complete solution. While XML+DTDs is indeed a very promising vehicle for structuring and transporting messages for business-to-business commerce, it is not itself a semantic theory of what those messages say. We need the semantic theory. Once we have that, XML can be used to embody it for applications.

Drawing on previous work, I will present the elements of my formal semantic theory for business messaging (the "lean events theory"). With examples from this theory before us, we can get a more proper view of the semantic transparency problem (aka: the spanning problem). This is not a problem that can be made to go away entirely, but we can live with it and do commerce.

1 EDI and the Transparency Problem

The standard EDI protocols (including the X12 and EDIFACT series) are of long-standing and effective use.¹ Yet, they have repeatedly been criticized for poor design, confusing or absent semantics, and much else.² Many of these criticisms are indeed on the mark. The main conclusions that the critics have drawn are also correct: business-to-business e-commerce is expensive and difficult to set up and maintain, because of shortcomings in the design concepts underlying standard EDI.

The problem is particularly acute for SMEs (small and medium-sized enterprises), which all too often find the cost of initial setup prohibitive. This cost includes substantial investment in legal services, management and staff time, and software. In the world of EDI, this is often referred to as the *first trade problem*. Once an EDI system is up and running smoothly, the incremental costs can be quite low and the incremental benefits very high. The problem is getting the *first* messages sent and working properly. That typically involves a large fixed cost.³

And, the first trade problem is hardly peculiar to the world of EDI. More broadly, it is recognized that there are important conceptual and technical issues to be addressed if the common vision—of millions of artificial agents cruising the Internet and doing deals for their owners in *ad hoc* and opportunistic ways—is to be realized. This is sometimes called the *spontaneity problem* [Dic00]; it is a more challenging generalization of the EDI first trade problem.

Something needs to be done to address these problems—first trade, spontaneity, etc.—but what?

A central issue—and the focus of this paper—is the fundamental question of *semantic transparency*: When S (speaker) sends A (addressee) a message, how does A's machine (or indeed A itself) know what the message is about, what it means? Or, given that A needs to understand a particular message, how can this be achieved at low cost and hence in a maximally-automated fashion?

Given proper standards, message meanings are determined and computers can be programmed to act appropriately to the intended message meanings. This is something that happens millions of times daily. The rap against EDI has been that proper standards have not in fact been created, and indeed cannot be made because of the misguided way in which the EDI standards are designed. The worry associated with the spontaneity problem is that we

¹ See, e.g., [Emm93, Emm94, Kim91] for valuable general introductions to electronic data interchange (EDI).

² For a sample of the criticisms, many from a basically friendly perspective, see: [AY96], [BLW97], [Dic00], [Kim91], [Kim99], [KM93a], [KT00], [Leh96], [Sal95], [Ste94], and [Ste96].

³ On the EDI first-trade problem see [Lee99, TT98a] and papers throughout [AY96].