



## New research directions for data and knowledge engineering: A philosophy of language approach

Andrew Basden <sup>a,\*</sup>, Heinz K. Klein <sup>b,c,✕</sup>

<sup>a</sup> Informatics Research Institute and Salford Business School, University of Salford, Manchester M5 4WT, United Kingdom

<sup>b</sup> School of Management, SUNY, Binghamton, USA

<sup>c</sup> Informatics Research Institute, University of Salford, Salford, United Kingdom

### ARTICLE INFO

#### Article history:

Available online 28 May 2008

This article is dedicated to the memory of Heinz Klein, a great, wide-ranging and prolific thinker, a wise and stimulating mentor, and a close friend, but who sadly passed away just before this paper went to press.

#### Keywords:

Text understanding  
Knowledge bases  
Models  
Ontologies  
Life-meaning  
Text-meaning  
Philosophy of language perspective  
Speech act theory  
Theory of communicative action  
Language action perspective  
Dooyeweerd's cosmonomic philosophy

### ABSTRACT

This article advances a strategic proposal that would enable future research in data/knowledge engineering and natural language processing to take a broader range of meanings into account than can be derived from analyzing text with current methods focusing on syntactic and semantic text meanings. It advocates drawing on the knowledge and common sense meanings that users acquire from NL interactions in their lifeworld, i.e. the 'life-meanings' that are created and shared in social communities. Three philosophical language perspectives are described to derive a research program for incorporating life-meaning based methods into contemporary DKE/NLP research.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

Suppose you live in Scotland or some 'cold' climate. "Were you born in a tunnel?" your mother asks as you enter the room. While you might be tempted to reply humorously, "No, in a hospital", you know what she means and you turn back and shut the door. Tunnels are draughty places (and in other parts of Britain 'tunnel' is 'barn' and in Sweden, 'church').

Ontologies and knowledge bases are often constructed from texts and attempt to model the meaning of the text. But this can be difficult when texts include such statements as this, because the meaning of the text is not immediately obvious from the words or phrases used (what we will call text-meaning), but must take account of how it is used in the lives of speaker or hearer (what we will call life-meaning).

\* Corresponding author.

E-mail address: [a.basden@salford.ac.uk](mailto:a.basden@salford.ac.uk) (A. Basden).

✕ Deceased.

This may be illustrated by the simple task of text retrieval. A user who wished to find texts that refer to the closing of doors would completely miss those that included “Were you born in a tunnel?” unless they possessed the life-knowledge that this idiom is sometimes used when “Please shut the door” is intended. Alternatively, a researcher might wish to determine the frequency of use of various idioms for “Please shut the door”, and again the life-meaning is essential to high quality text retrieval. However, once the idiom has become widespread in society, then the text retrieval search engine itself might learn that “born in a tunnel” refers to door-closing.

This kind of semantic link lies at the root of Berners-Lee’s [6, p. 169] vision for the Web:

“Machines become capable of analysing all the data on the Web – the content, links and transactions between people and computers. A ‘Semantic Web’, which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines, leaving humans to provide the inspiration and intuition. The intelligent ‘agents’ people have touted for ages will finally materialise”.

But will they? Should they? Winograd [64, p. 11] warns that such a vision is not only difficult but likely to be inappropriate, and suggests,

“a key design issue: let people do the interpretation of natural language, and let the program deal with explicit declarations of structure”.

Life-meaning is one reason why Berners-Lee’s vision might be inappropriate<sup>1</sup> and provides support for Winograd’s view. Life-meaning can affect at least the following tasks in data and knowledge engineering (DKE) that involve interpretation of language (spoken or written):

Knowledge base construction – If the knowledge engineer misunderstands what the domain expert says, the resulting knowledge base will malfunction, or at least mislead its users.

Ontology construction – It is important to find the exceptions as well as the average knowledge if the ontology is to be generally applicable, and many exceptions relate to human life.

Intelligent agents – These should behave as human agents would in similar situations.

Virtual worlds – They need to be believable in all respects.

Model checking and comparison – The tacit life-assumptions inscribed into each model can modify its intended meaning in major ways.

Information systems development (ISD) – Misunderstandings between members of the ISD team can arise from different assumed life-meanings.

Life-meaning can also affect the following tasks in natural language processing (NLP):

Text understanding – Models and ontologies automatically derived from texts will be deficient unless the full meaning of texts is taken into account.

Text summarization, indexing and correction – Unless life-meaning is taken into account, summaries and indices might be misleading, and wrong corrections might occur.

Text generation – Text generated from models tends to be bland and boring if it states explicitly what is taken for granted by its readers and does not employ devices like humour.

Machine translation – see the example in Section 2!

If these tasks are to be applied in non-technical applications or those outside the professional arena, such as social networking, virtual worlds or computer games, then life-meaning will be crucial to their success; as we shall see below, it can also be important in technical applications. If this list is at least partially correct, then the potential benefits of more powerful, computer-supported, partially automated NLP could be huge. We shall reflect on this point again in the conclusions.

Before the benefits can be reaped, we need to admit that it is by no means simple to explain how life-meaning could be taken into account. Understanding life-meaning in general is a prerequisite for addressing the more technical types of life-meaning encountered in some of the tasks above. In particular, it is now generally accepted that it is impossible to capture the entire meaning in rules or other structures expressed in a knowledge representation language. Much of the life-meaning, especially of the user, will be applied during the very act of interpreting the information presented by the computer system. In such cases, it is not necessary to encode the life-meaning in rules or structures within the computer system, supporting Winograd’s view above. Regardless of whether life-meaning is encoded or is left to the user, it is important for researchers to understand how the life-meaning of texts (in the widest sense) affects human understanding and machine interpretations. It therefore needs to be made as explicit as possible. Very little research in DKE or NLP has tackled this and therefore it is an appropriate topic for this article.

Text analysis research is strongly influenced by what Winograd [64] calls the traditional perspective on language, which begins with syntax, ventures into semantics when this fails, and only then, and grudgingly, takes pragmatics into account. This order is reversed for the Language/Action Perspective (LAP): first pragmatics, then semantics and lastly syntax. Lyytinen [37] briefly reviewed five major perspectives on language and argued that the perspective taken has major impacts on infor-

<sup>1</sup> Berners-Lee seems to express opinions similar to Winograd’s shortly afterwards, but then returns to the optimistic view of machine processing.

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات