PRELUDE

The human personality acts or exists through mental processes, actions and communication. By their very nature these manifestations are discrete, localized and transient. Equally, the success of the attempts by the personality to ‘know itself’ can only be halting and sporadic. It is no surprise that there is a constant gap between the assertion and its reception – whether by oneself or by others. This perhaps is the metaphysical genesis of ambiguity.

But it is not given to us to live with ambiguity especially when we are unreasonably certain that a wholesome meaning, an immaculate significance underlies the thoughts, actions and communication. It is therefore a human need to overcome or perhaps to resolve ambiguities. The ambiguity is not of one type. The various types have different effects.

Fortunately, a context is always available. The personality of the actor, the daily routines, current culture, historical traditions, the grammar and nature of language are some of the significant factors which provide the context.

The context reveals the ambiguity and it is in that context that we unravel the ambiguity. Ambiguity does not walk by itself. When the context appears inadequate or appears to point in a different direction, we need to either broaden the context or provide rules for interpretation of the context.

Ambiguity is one of the qualities of personality or a category of culture. We come to expect ambiguities and not even notice their presence. Even when they are obviously present, we may pass over with effortless skill. But it is a different story for a machine. When a machine is set the task of finding the meaning or translating a passage, it will not have the advantage of having developed through personal and societal interactions and relationships. Rules will have to be provided to the machine not only for resolving ambiguities but also to identify or locate ambiguities. An ambiguity may appear trivial in the context of experience. But for a machine every ambiguity is a challenge and nothing is trivial. The machine will have to be taught even about barely perceptible ambiguities and their possible resolutions.

This thesis is inherently experiential. In order to deal with ambiguous entities which we come across during the process of machine translation, it would be necessary to draw from a wide range of contextual data, which can be located in huge corpora along with different NLP (natural language processing) tools like POS-taggers. Thus, the accessibility to corpora and development of NLP tools are conditions for undertaking work of the type chronicled herein.

It is my hope as a linguist that these new insights into procedures of developing techniques for word sense disambiguation will enable us to evolve applications that better support both Computational Linguistics and science of Artificial Intelligence. Indeed, by increasing our understanding of the complex phenomenon of linguistic ambiguity, we can better inform the design of advanced NLP tools and applications such as text to speech and speech to text conversions.