

# Modelling an Author's Style and Reaching Quickly to Stylometric Conclusions by Using Multi-Valued/Fuzzy Logic

By Dimitrios Foukis, a PhD research student at the Faculty of Arts, Design and Humanities, De Montfort University.

Funded by DMU, Email: [p1123303x@my365.dmu.a.uk](mailto:p1123303x@my365.dmu.a.uk), Supervised by Professor Gabriel Egan.

## Introduction

In the traditional logic we say that an object ( $x$ ) belongs ( $\exists$ ) to a set or collection of objects ( $S$ ) or does not belong ( $\nexists$ ). In Fuzzy Logic, there are truth values and the concept of **membership degree**  $\mu_S(x)$  of an element  $x$  to a class/set ( $S$ ) in the range of 0.0 to

1. As the grade of membership approaches the unity (1), the higher is the membership. In this context, by applying Fuzzy Logic in the stylometric area, the goal is not to classify an anonymous, disputed play in the limited range of binary values as Shakespearean (1) or not Shakespearean (0), but to attempt to find with 'fuzzy' accuracy how much Shakespearean indeed this play is ('evidential interval' of degrees of Shakespearean Indices of Similarity [S.i.S]).

## Research Objectives

- To show how we can apply by analogy the principles of Fuzzy Logic in the micro-engineering and stylometric area as well.
- To model stylistic traits in a way (Fuzzy-logic) that has never been tried by anyone else in the area of computational stylistics. One model for each genre (comedy, tragedy, history).
- To employ counts of semantic sets of words as principal stylistic markers and to turn problems of authorship attribution into problems of automated approximate reasoning.

## Methodology

- This investigation falls into the field of the positivist research paradigm, which aims at the 'fuzzified-defuzzified' assessment of an author's quantifiable stylistic traits.
- Triptych: The **propositional calculus** with the use of a **Boolean algebraic operations** (max =  $\cup$  = 'or', min =  $\cap$  = 'and', 'minimum Shakespeareanness' = threshold) and the **Set Theory**.

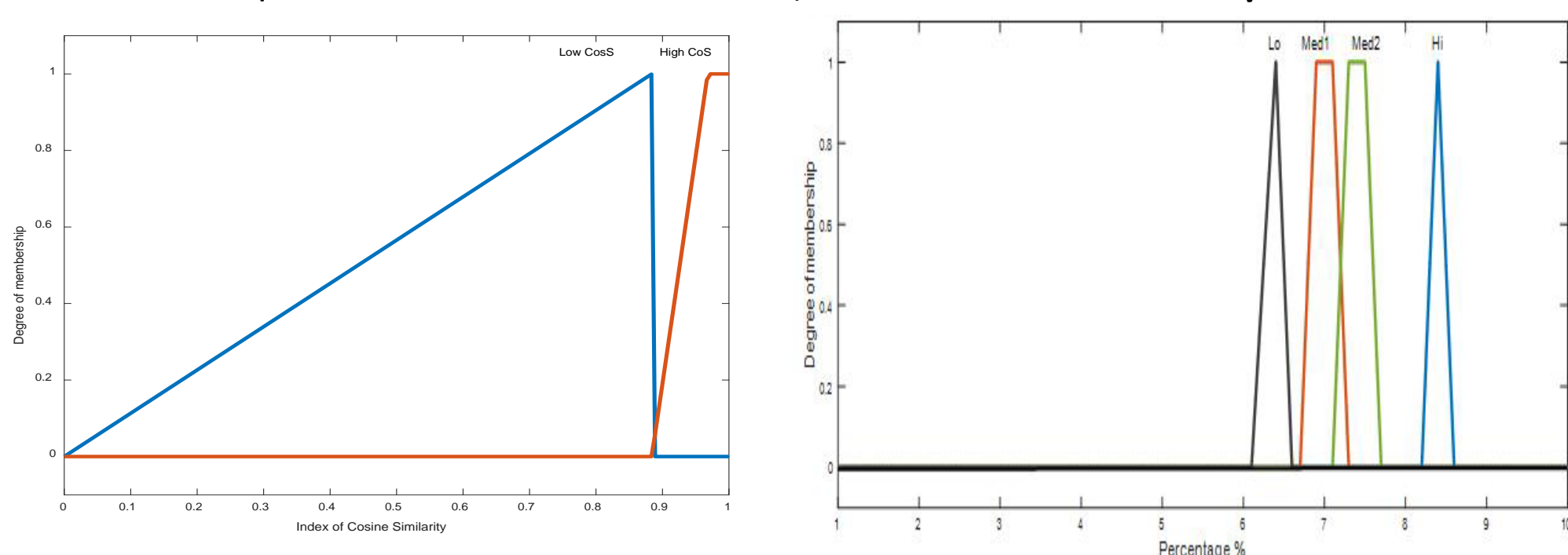


Figure 1: Membership functions (MF)-classes of the Index of Shakespearean Cosine Similarity. Figure 2: A Set's Membership functions (MF)-classes of Shakespeareanness (based on the counts of a Set of 20 words' frequencies ÷ whole extract of text (≈ 10,000 words)).

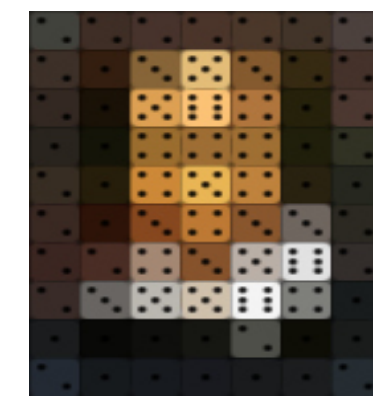


Figure 3: 'Profiling'

- ✓ **Resultant Stylistic Classifier** through Computer Simulation
- ✓ **7 Stylistic Markers = 7 Inputs (I) →**  
→ min =  $\cap(\mu I1 \text{ to } \mu I7)$  → Rules → 2 S.I.S Outputs

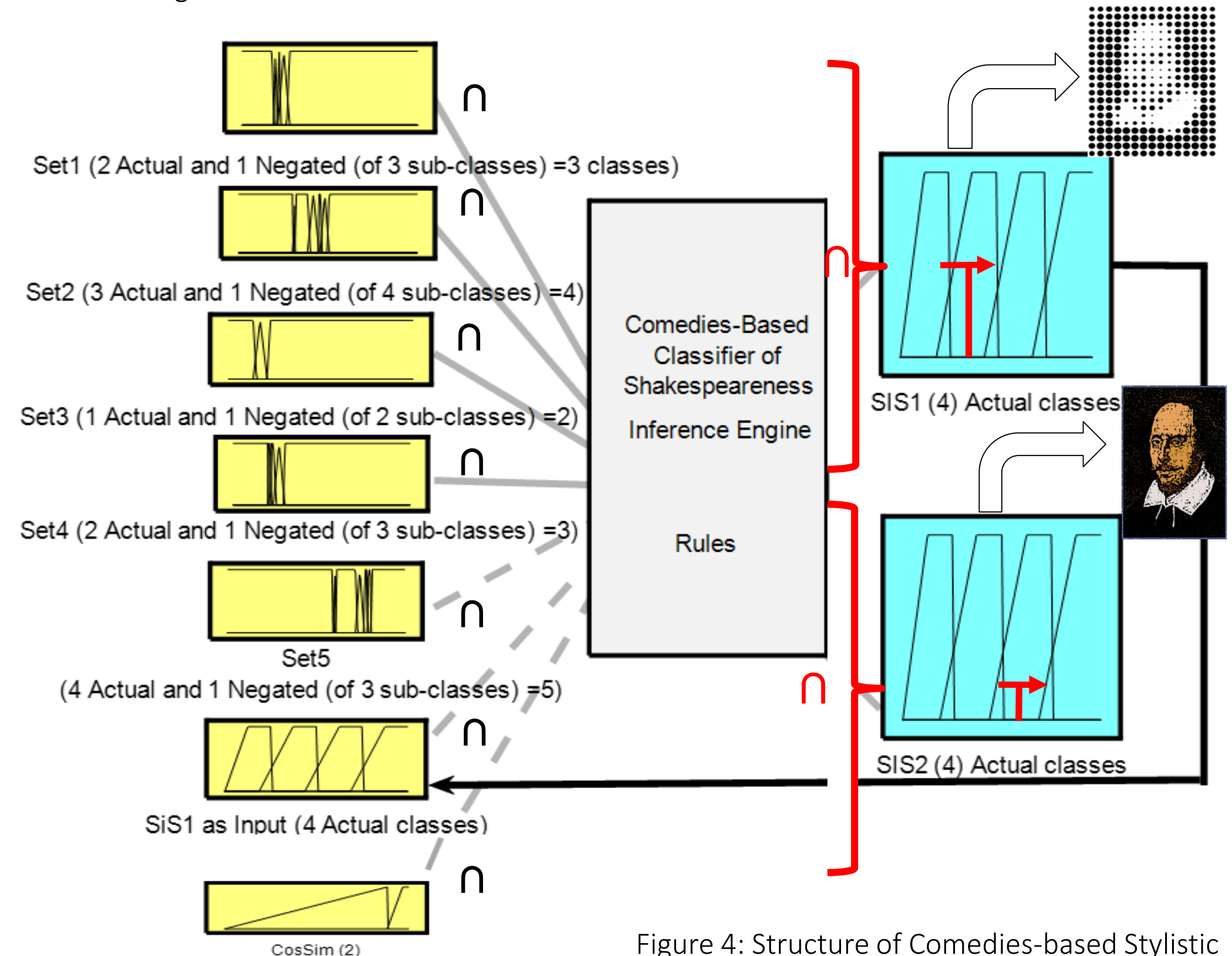


Figure 4: Structure of Comedies-based Stylistic Classifier-Simulator

## Conclusions

- Yes, 'Shakespearean' counts of stylistics features, such as an index of counts of words' frequencies (cosine similarity) between two vectors-documents and the counts of sets of words, of sole-authored, well attributed Shakespearean plays can be efficiently modelled as 'Shakespearean' triangular or trapezoidal classes in a two-dimensional coordinate system (see Figures 1 and 2).
- Yes, it is possible by employing counts of sets of words' frequencies and Fuzzy Logic to investigate the authorship of disputed, anonymous plays of the Early Modern English Period.



Figure 5: Authorial Pattern Recognition

- Yes, it is possible to pave the way for a new automated Fuzzy-Logic-based universal stylistic classifier of any kind of written text and for various purposes.

## Bibliography

- Craig, Hugh, and Arthur F. Kinney, eds. 2009. *Shakespeare, Computers, and the Mystery of Authorship*. Cambridge: University Press.
- Sivanandam, S.N., S. N. Deepa, and S. Sumathi. 2007. *Introduction to Fuzzy Logic Using Matlab*. Berlin: Springer.
- Zadeh, L. 1968. "Communication. Fuzzy Algorithms." *Information and Control* 12 (2):94–102.
- Zadeh, L. A. 1965. "Fuzzy Sets." *Information and Control* 8 (338-353).