Imaging and Information Retrieval: Variations on a Theme*

F. Crestani[1,2], F. Sebastiani[3] and C.J. van Rijsbergen[2]

[1] Università degli Studi di Padova, Padova, Italy

[2] University of Glasgow, Glasgow, United Kingdom

[3] Consiglio Nazionale delle Ricerche, Pisa, Italy

Abstract

(Standard) Imaging is a method for the revision of probability functions originally proposed in the philosophy of language as a semantics for conditional logic. Recently, Standard Imaging and a variant of it, called General Imaging, have successfully been applied to the estimation of the probability of relevance in Information Retrieval (IR) by Crestani and van Rijsbergen. The experimental results they have obtained show that these methods perform better than a number of more established approaches, such as retrieval by joint or conditional probability.

In this paper we report preliminary results on three main generalisations of these methods and their application to IR. These generalisations are orthogonal (and they may thus be freely combined), as they address three orthogonal issues in the probability kinematics of Imaging.

The first generalisation, that we call *Proportional Imaging*, is a variation of General Imaging that is better suited to those cases in which similarity between "possible worlds" has a quantitative nature; this is indeed the case in the application of Imaging methods to IR, where keywords play the role of possible worlds. The idea that underlies Proportional Imaging is that the probability of an $\overline{A}$-world $\overline{w}$ should be distributed to *all* A-worlds w_i in a way that is proportional to the degree of similarity between $\overline{w}$ and the w_i's.

*This work has been carried out in the context of the project FERMI 8134 - "Formalisation and Experimentation in the Retrieval of Multimedia Information", funded by the European Community under the ESPRIT Basic Research scheme.

The second generalisation, that we call *Jeffrey Imaging*, is an extension of Imaging methods that addresses the revision of probability functions on the basis of *uncertain* evidence; in IR terms, this can be applied to the computation of a sophisticated form of relevance feedback, namely one that can account for user relevance judgments in the form of "degrees of relevance". The idea that underlies Jeffrey Imaging is that only *part* of the probability of a $\overline{A}$-world $\overline{w}$ is redistributed to the A-worlds, similarly to what happens in the possible-worlds view of Jeffrey Conditionalisation. The probability that $\overline{A}$-worlds retain has the effect of allowing the revised probability of A to be nonzero, as desired.

The third generalisation, that we call *Mixed Imaging*, consists of a method for the revision of probability functions that combines standard Bayesian revision with Imaging-type revision. This is based on the intuition that, when similarity between a $\overline{A}$-world $\overline{w}$ and all A-worlds w_i tends to 0, similarity seems hardly intuitive as a criterion on which revision should be performed; worse, when the similarity between $\overline{w}$ and w_i is 0 for all i, Imaging methods are undefined. Bayesian conditionalisation seems then the right choice in these cases. The idea that underlies Mixed Imaging is thus that the probability $\mu(\overline{w})$ of an $\overline{A}$-world $\overline{w}$ is redistributed to A-worlds only partly by an Imaging method, while the rest is redistributed to A-worlds according to Bayesian conditionalisation. Exactly how much of it goes by Imaging is determined by a parameter α that varies according to the "base" variant of Imaging (Standard, General, Proportional) chosen; for example, in the case of Standard Imaging, α is the degree of similarity between $\overline{w}$ and its most similar A-world.

We also report on some preliminary negative results concerning *Weighted Imaging*, a would-be extension of Imaging methods that should account for those cases in which formulae are true at possible worlds only to a certain degree; in IR terms, this would be particularly suited to addressing the case in which documents and information needs are represented by vectors of *weighted* keywords. Unfortunately, we have proven that in the weighted case it is impossible to give revisions of probability functions guaranteeing that $P_A(A) = 1$.