**Department of Decision Sciences - Bocconi University**
Via Roentgen 1 - 20136 Milano
Tel. 02-58365632 - Fax 02-58365630

# SEMINAR

## "Probabilistic expert systems for DNA Mixtures"

## Julia Mortera
(Università Roma Tre, Italy)

## Thursday, 18th December 2008 – h. 15.30
## Room N17 – Piazza Sraffa 13 - 20136 Milano

---

### Abstract:

A new methodology, based upon probabilistic expert systems, for analysing forensic identification problems involving DNA mixture traces using quantitative peak area information is illustrated.

The peak area is modelled with conditional Gamma distributions. The probabilistic expert system can be used for ascertaining whether individuals, whose profiles have been measured, have contributed to the mixture, but also to predict DNA profiles of unknown contributors by separating the mixture into its individual components. The potential of our probabilistic methodology is illustrated on case data examples and compared with alternative approaches. The advantages are that identification and separation issues can be handled in a unified way within a single probabilistic model and the uncertainty associated with the analysis is quantified.

The methodology yields a coherent way of predicting genotypes of unknown contributors and assessing evidence for particular individuals having contributed to the mixture. The advantages of a probabilistic model-based approach are that there is a natural and directly interpretable quantification of all uncertainties associated with the analysis; in particular, the posterior distribution of the mixture proportion can be computed.

Furthermore, the system itself would warn against trusting an uncertain prediction, by yielding an associated low classification probability.

In addition, the analysis is extendable to similar but different situations using the modularity and flexibility of the probabilistic expert system approach. The same approach can be used to deal with artifacts such as silent alleles, stutter peaks and dropout alleles. It can also be used for analysing two-person and three-person mixtures that combine peak area information from multiple independent samples. Also possible kinship relationships between contributors to mixtures can be handled in this framework.