

Philippe ABBAL
INRA UMR SPO

Responsable de la cellule représentation des connaissances à l'UMR SPO.
Thèse en cours sur le sujet : analyse et modélisation de la qualité d'un vin.
IR1, Elève à l'école doctorale SP/SA SupAgro

- A/ Problématique actuelle à l'UMR SPO et contexte
- B/ Deux approches liées à la qualité du vin

- 1- *Prédire le potentiel qualitatif d'un vignoble*



- 2- *Modéliser la qualité d'un vin*



A/ Contexte

- Actuellement il n'existe pas d'outil permettant d'intégrer toutes les données disponibles dans le processus d'élaboration du vin afin de prédire sa qualité. Ce travail une fois mené a terme et validé constitue une avancée dans le domaine .
- S'inscrit dans le cadre de la cellule représentation des connaissances de l'UMR.
- Sujet multidisciplinaire : mathématiques, connaissances en viticulture et œnologie, analyse sensorielle, statistiques, informatique (développement d'applications et bases de données),
- Collaborations: A. Carbonneau, UE P.R., Eric Matzner (Ensa de Rennes), Paul Coulon (Domaine de Bearenard), Pascal Neveu (Mistea).
- Axe fortement encouragé par le département Cepia

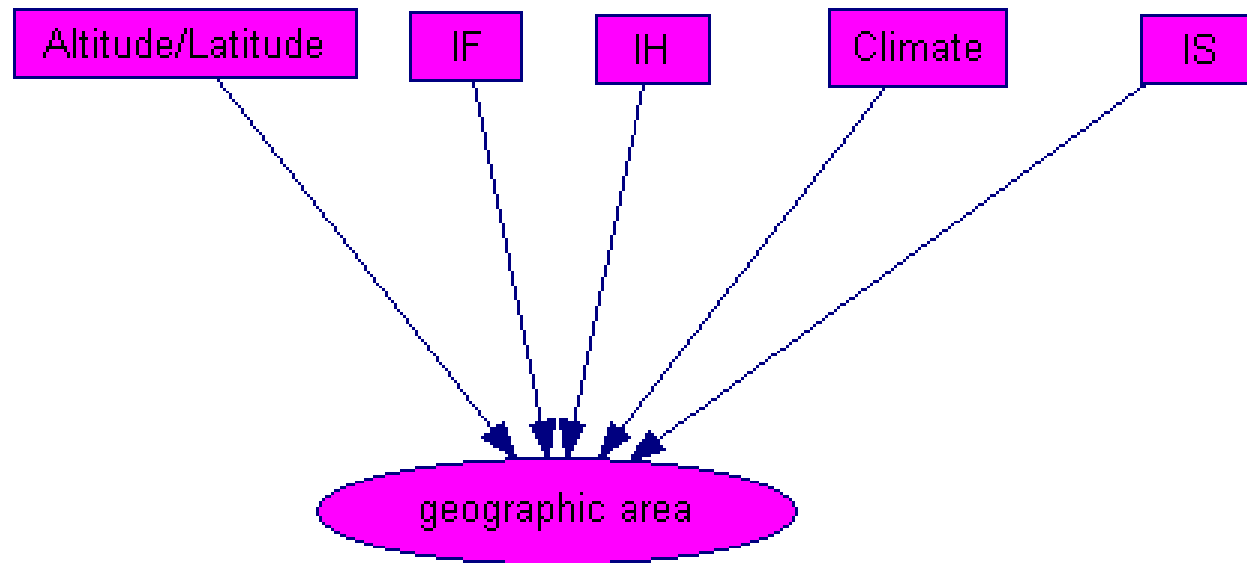
B1 / Première partie

- Comment prédire le potentiel qualitatif d'un vignoble ?
- Développer un modèle particulier assurant la synthèse des connaissances dans ce domaine
 - Quelle méthodologie choisir ?
 - Réseaux de neurones, logique floue , réseaux bayésiens ?

Choix retenu : le réseau bayésien

- Basé sur une ***représentation graphique*** du problème posé et la **dépendance** de variables entre elles
- Intègre variables quantitatives ou qualitatives
- Définis par des nœuds de décision, des nœuds aléatoires et des relations de dépendance (inférence)
- Modèle probabiliste

Exemple : schéma de principe de définition graphique d'un des nœuds du modèle



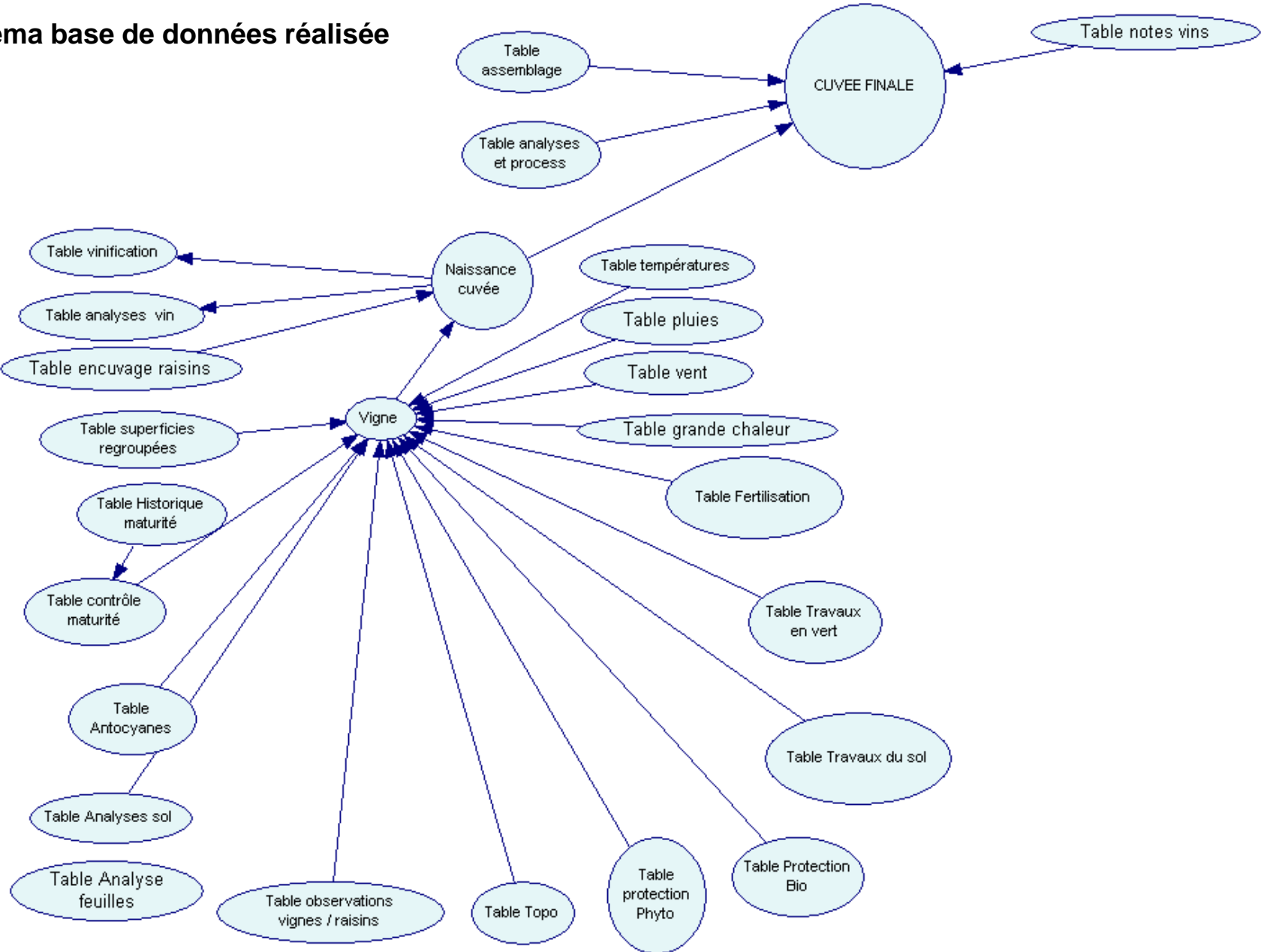
Traitements et calculs

- Basés sur
 - - 1/ un logiciel adapté aux réseaux bayesiens Genie 2 (Université de Pittsburg USA)
 - - 2/ un outil de pilotage et d'interface développé spécifiquement C#(Microsoft)
 - permet de saisir en les contrôlant les variables des nœuds de décision
 - génère le calcul des nœuds aléatoires en pilotant directement le cœur du logiciel bayésien
 - Intervient directement dans le moteur d'inférence SMILE (Structural Modeling, Inference, and Learning Engine)

B2/ Modéliser la qualité d'un vin

- Idée de départ : intégrer un ensemble de données quantitatives, qualitatives et appréciations d'experts afin de déterminer un modèle mathématique de prédiction de la qualité d'un vin.
- Origine des données : importante base de données viticoles, environnementales et œnologiques, observées pendant plusieurs dizaines d'années et dans des situations variées.
- Validation du modèle: variable de sortie à prédire : notes basées sur plusieurs sources : Robert Parker, notes données par les vignerons eux-mêmes, notes de jurys de concours.
- Validations sur UE INRA Pech Rouge.

Schéma base de données réalisée



Méthode utilisée

- Création d'un modèle d'expert (comportant 6 sous modèles) du type régression multilinéaire
- Analyse des données et validation des régressions linéaires multiples avec R (lm)
- Utilisation de méthodes de régressions non paramétriques à splines et noyaux (package ibr)

Perspectives et développements

- Doter notre UMR d'outils de représentation de connaissances liées au vin.
- Utiliser ces approches méthodologiques dans le cadre de notre travail de recherches en relation avec l'unité expérimentale de P.R.
- Poursuivre la validation du modèle en incluant l'analyse des cinétiques de fermentation et les outils du cati CODEX.

ANNEXES

- **GeNle**

- The GeNle (Graphical Network Interface) software package can be used to create decision theoretic models intuitively using the graphical click-and-drop interface. GeNle is a graphical interface with a fully portable Bayesian inference engine developed by the Decision Systems Laboratory and thoroughly tested in the field since 1998.

- **Primary Features:**

- Graphical editor to create and modify network models.
- Uses the SMILE Engine. You may develop models in GeNle and create a custom interface for them using SMILE.
- Open multiple networks and cut and paste sections of models between them.
- Complete integration with MS Excel, cut and paste data into internal spreadsheet view of GeNle.
- Cross compatibility with other software. Supports all major file types (e.g. Hugin, Netica, Ergo).
- Support for handling observation costs of nodes.
- Support for diagnostic case management.

- **SMILE**
- SMILE (Structural Modeling, Inference, and Learning Engine) is a fully platform independent library of C++ classes implementing graphical probabilistic and decision-theoretic models, such as Bayesian networks, influence diagrams, and structural equation models. Its individual classes, defined in SMILE API (Application Programming Interface), allows you create, edit, save, and load graphical models, and use them for probabilistic reasoning and decision making under uncertainty.
- SMILE supports directly object-oriented methodology. SMILE is implemented in C++ in a platform independent fashion. Individual classes of SMILE are accessible from C++ or (as functions) from C programming languages. As most implementations of programming languages define a C interface, this make SMILE accessible from practically any language on any system. SMILE can be embedded in programs that use graphical probabilistic models as their reasoning engines. Models developed in GeNIe can be equipped with a user interface which utilizes SMILE as the backend engine. SMILE is released as a dynamic link library (DLL). There are also several SMILE wrappers, such as SMILE.NET (.NET interface), SMILEX (Active X), jSMILE (Java interface), etc.