De l'identification de structure de réseaux bayésiens à la reconnaissance de formes à partir d'informations complètes ou incomplètes

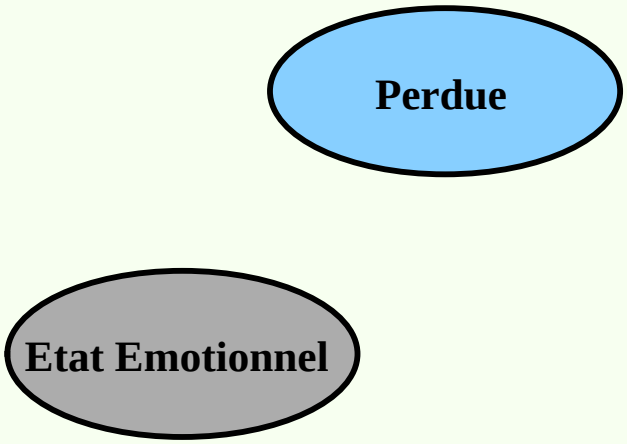
Olivier FRANÇOIS

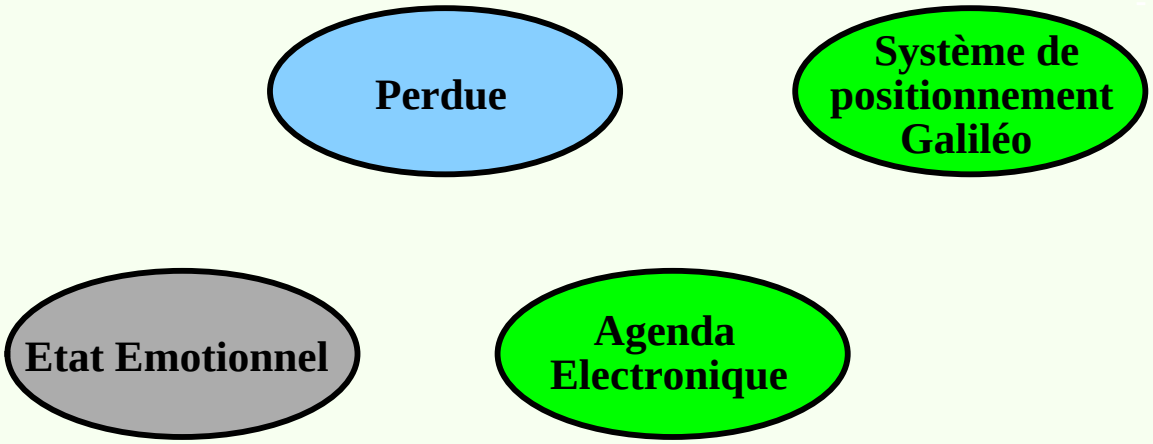
Encadrant : Philippe LERAY
Directeur : Stéphane CANU

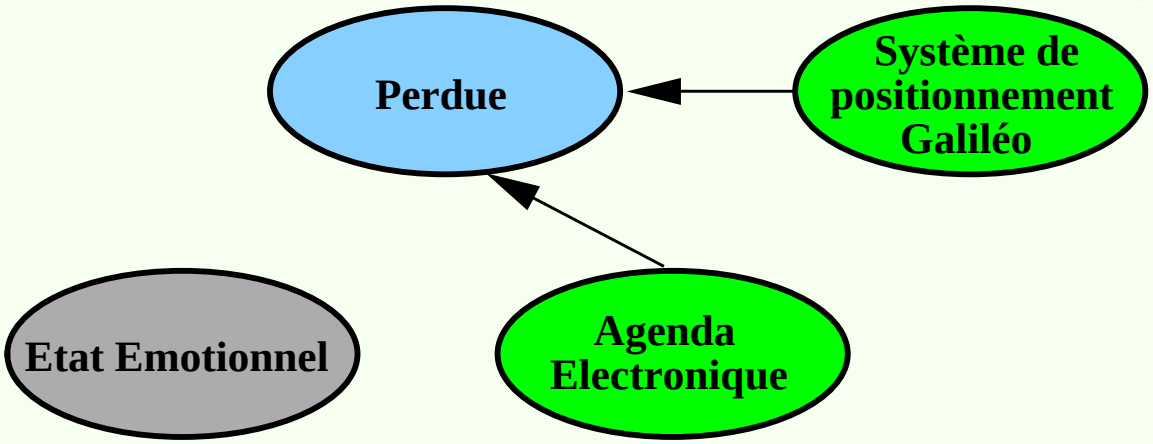
Laboratoire **LITIS** , Rouen

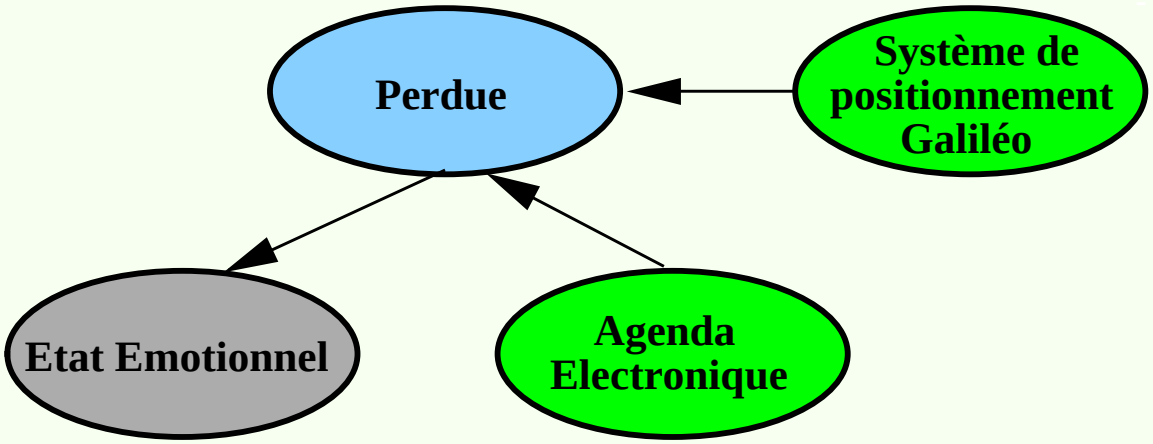
présenté le mardi 28 novembre 2006

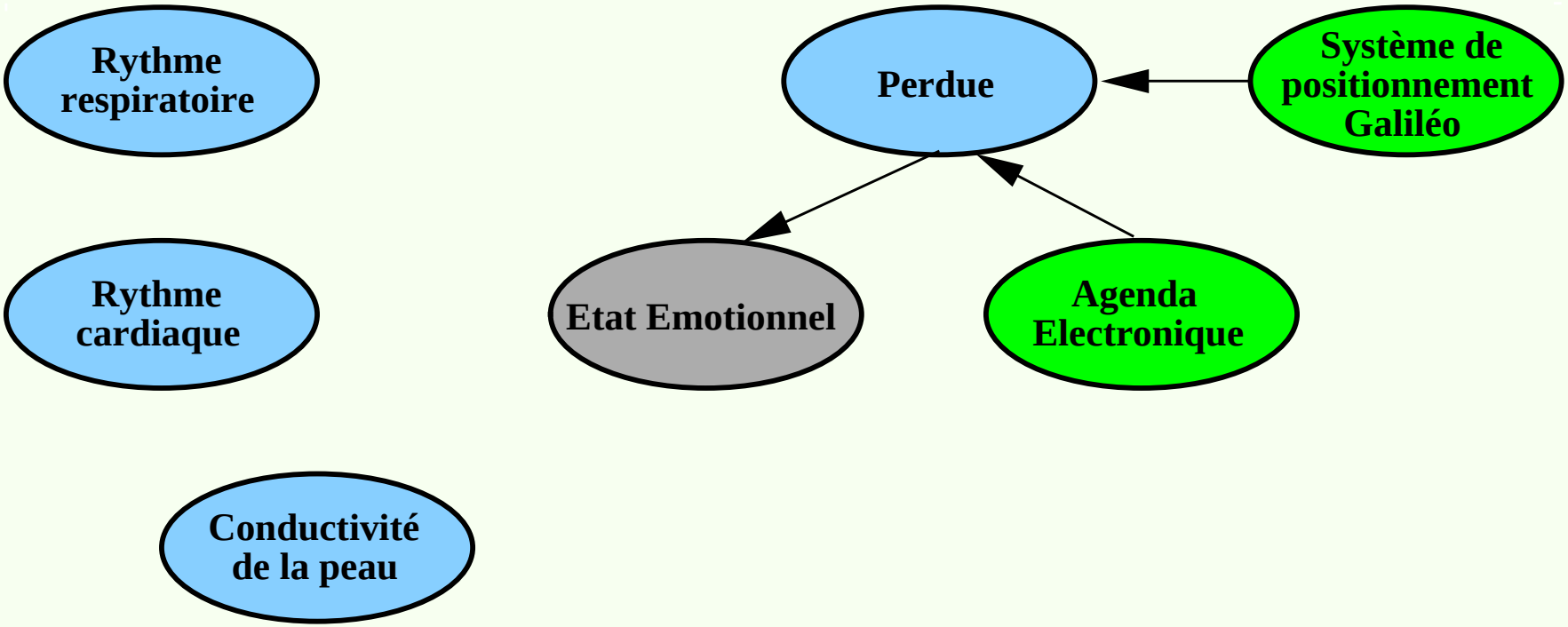
Etat Emotionnel

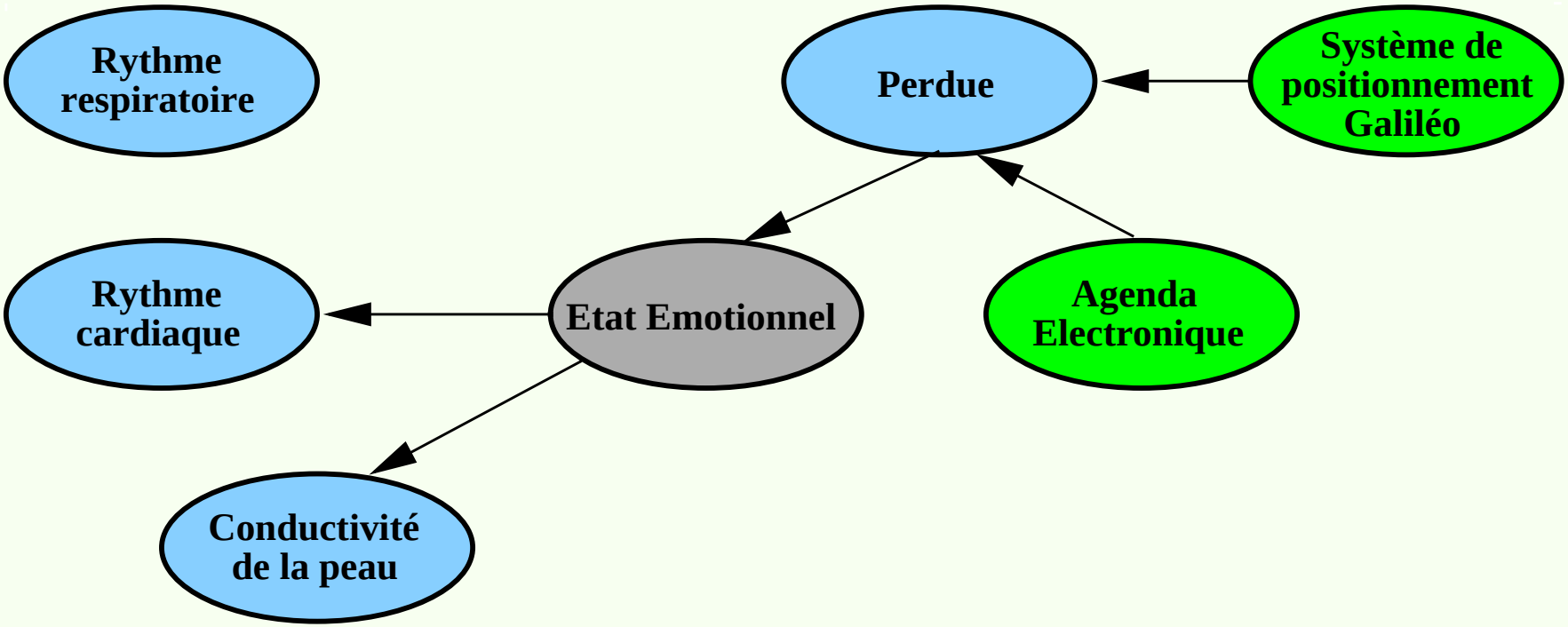

Perdue
Etat Emotionnel

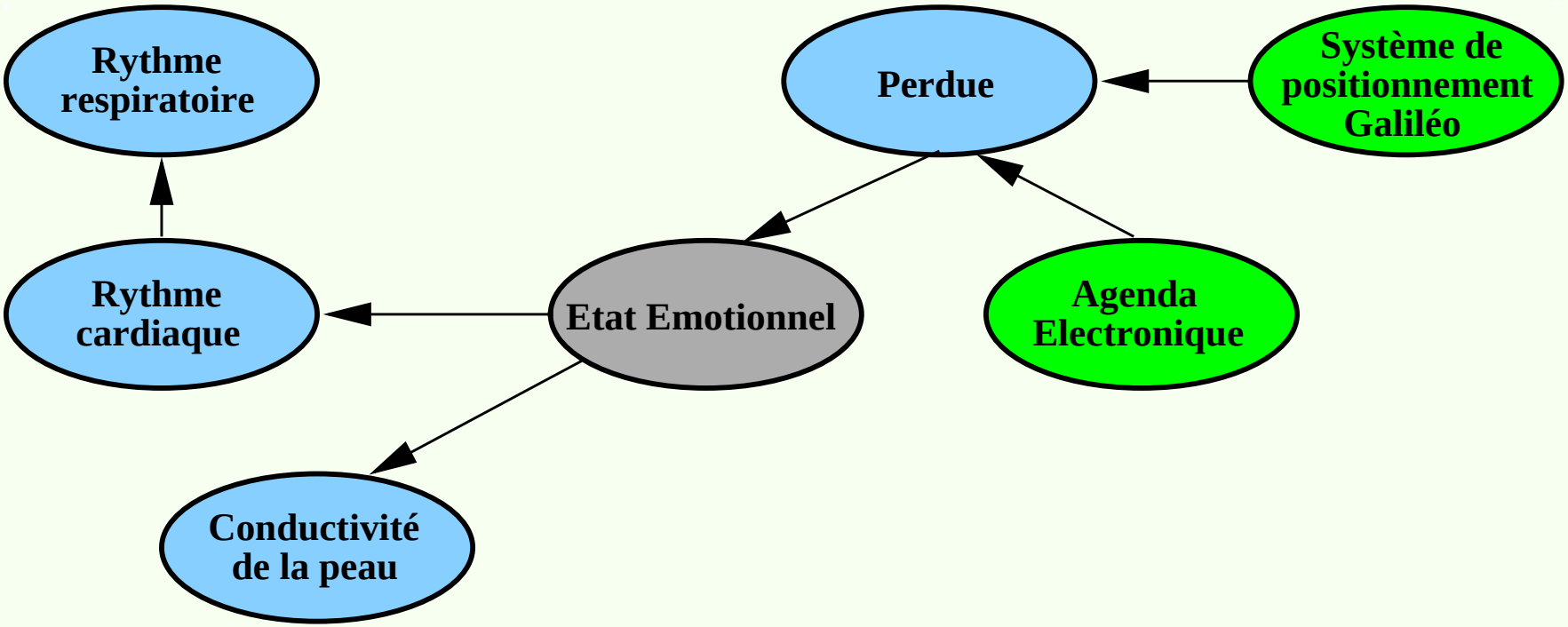

Perdue
Système de positionnement Galiléo
Etat Emotionnel
Agenda Electronique

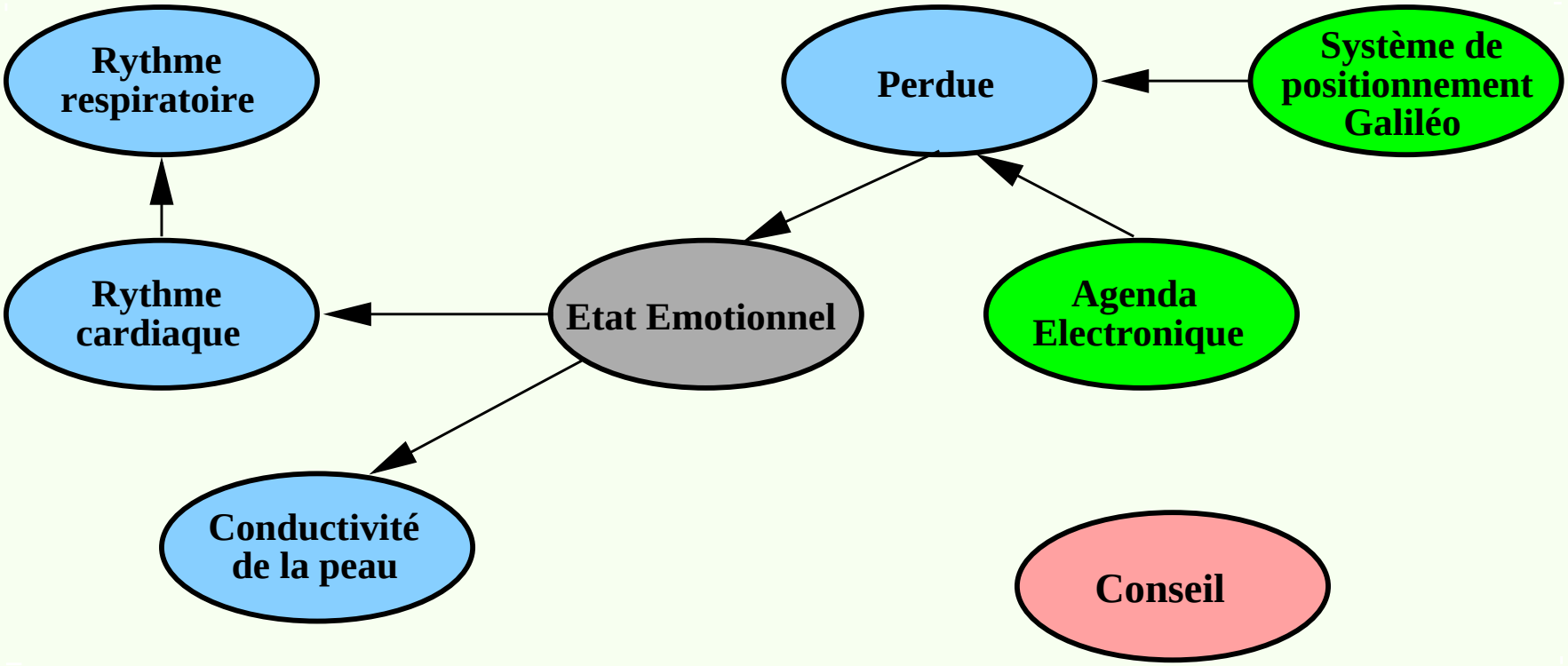

Perdue
Système de positionnement Galiléo
Etat Emotionnel
Agenda Electronique

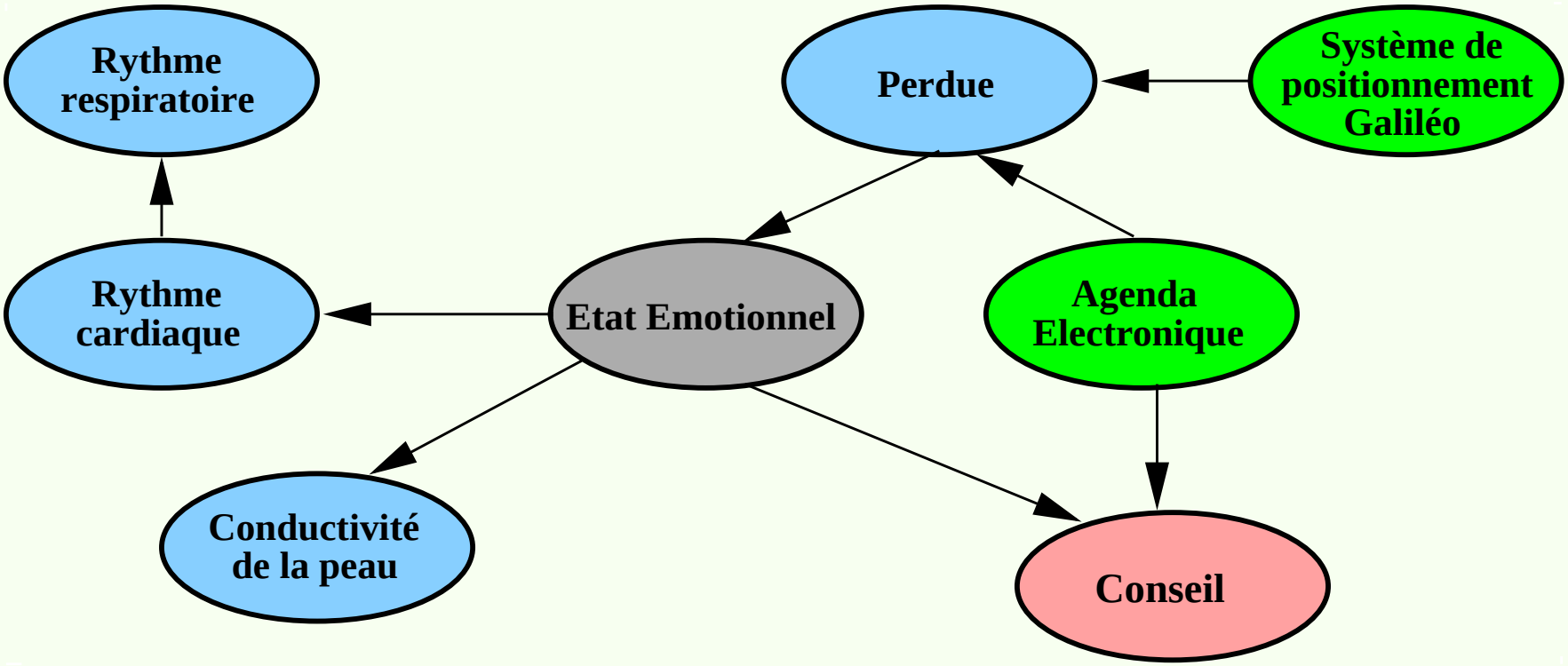

Perdue
Système de positionnement Galiléo
Etat Emotionnel
Agenda Electronique


Rythme respiratoire
Rythme cardiaque
Conductivité de la peau
Perdue
Système de positionnement Galiléo
Etat Emotionnel
Agenda Electronique

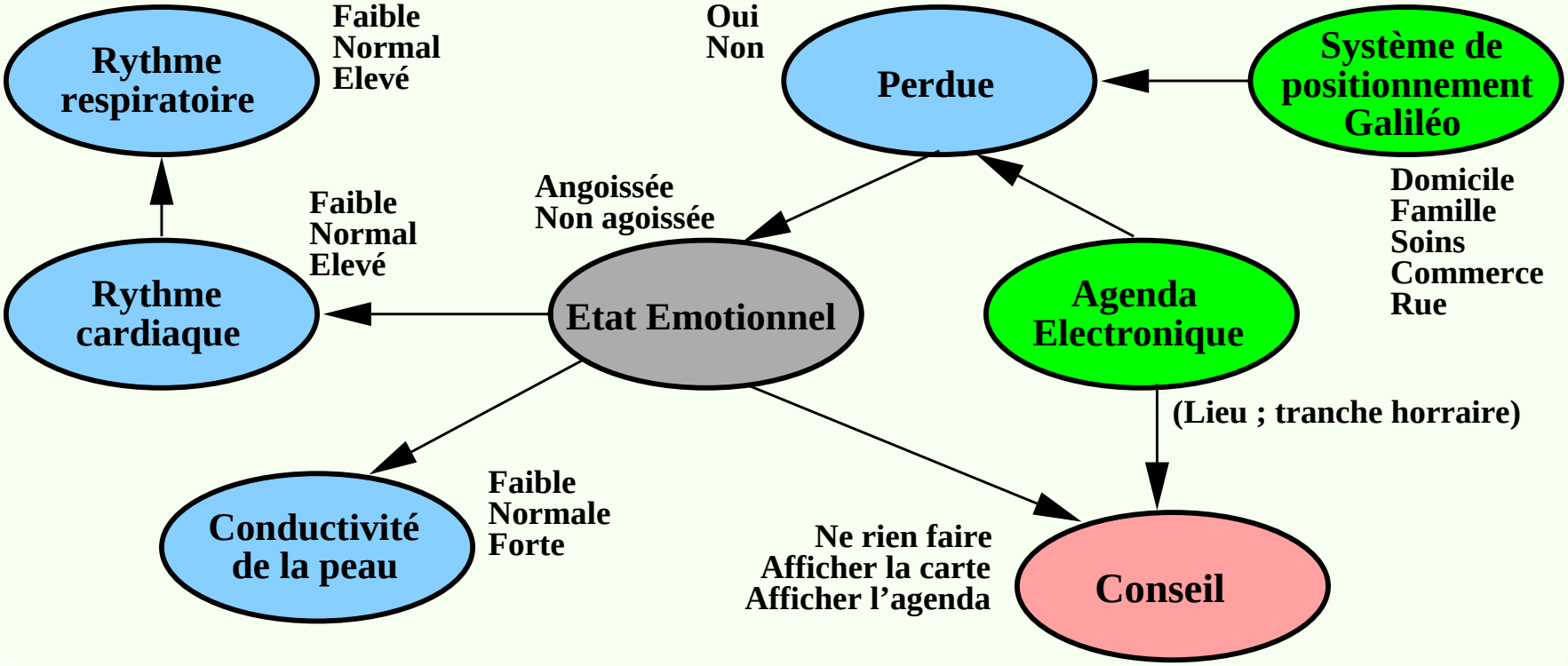

Rythme respiratoire
Perdue
Système de positionnement Galiléo
Rythme cardiaque
Etat Emotionnel
Agenda Electronique
Conductivité de la peau


Rythme respiratoire
Perdue
Système de positionnement Galiléo
Rythme cardiaque
Etat Emotionnel
Agenda Electronique
Conductivité de la peau

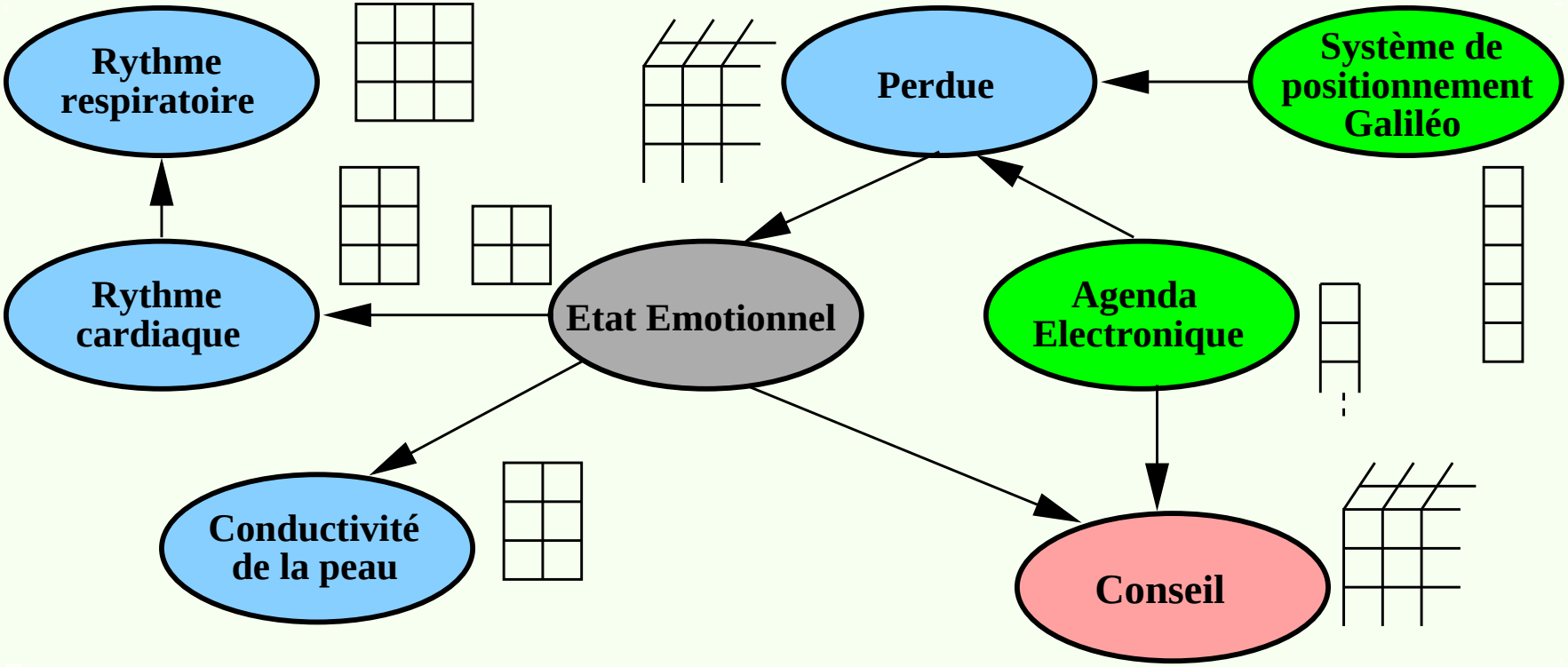

Rythme respiratoire
Perdue
Système de positionnement Galiléo
Rythme cardiaque
Etat Emotionnel
Agenda Electronique
Conductivité de la peau
Conseil


Rythme respiratoire
Perdue
Système de positionnement Galiléo
Rythme cardiaque
Etat Emotionnel
Agenda Electronique
Conductivité de la peau
Conseil

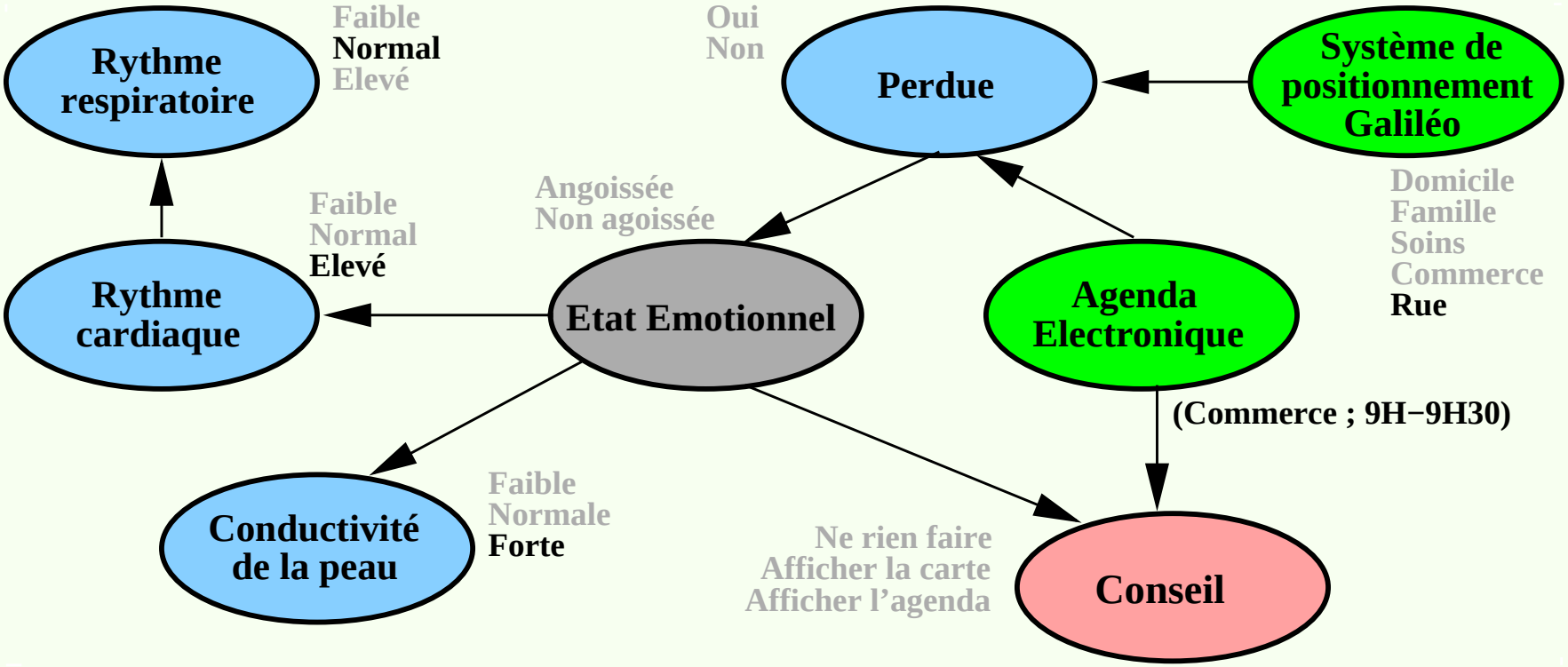
Définition d'un Réseau Bayésien



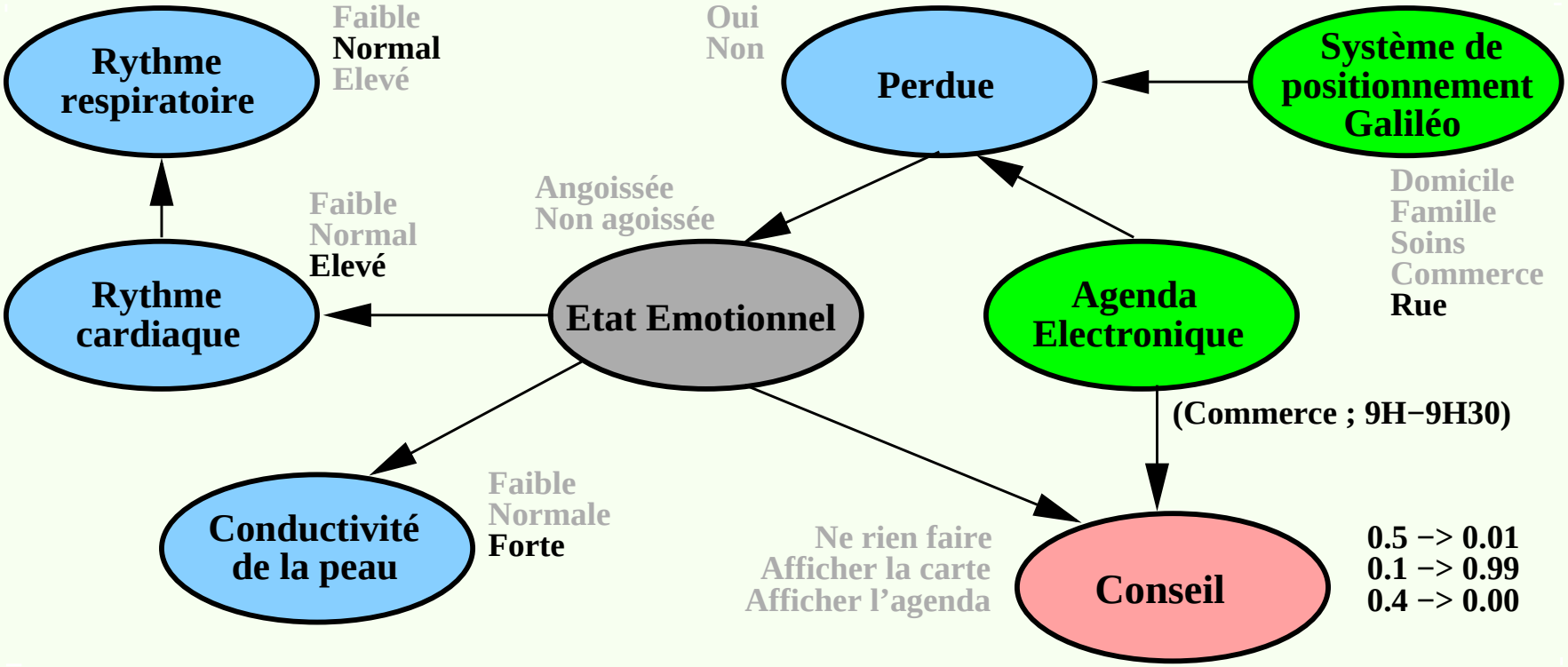
Définition d'un Réseau Bayésien



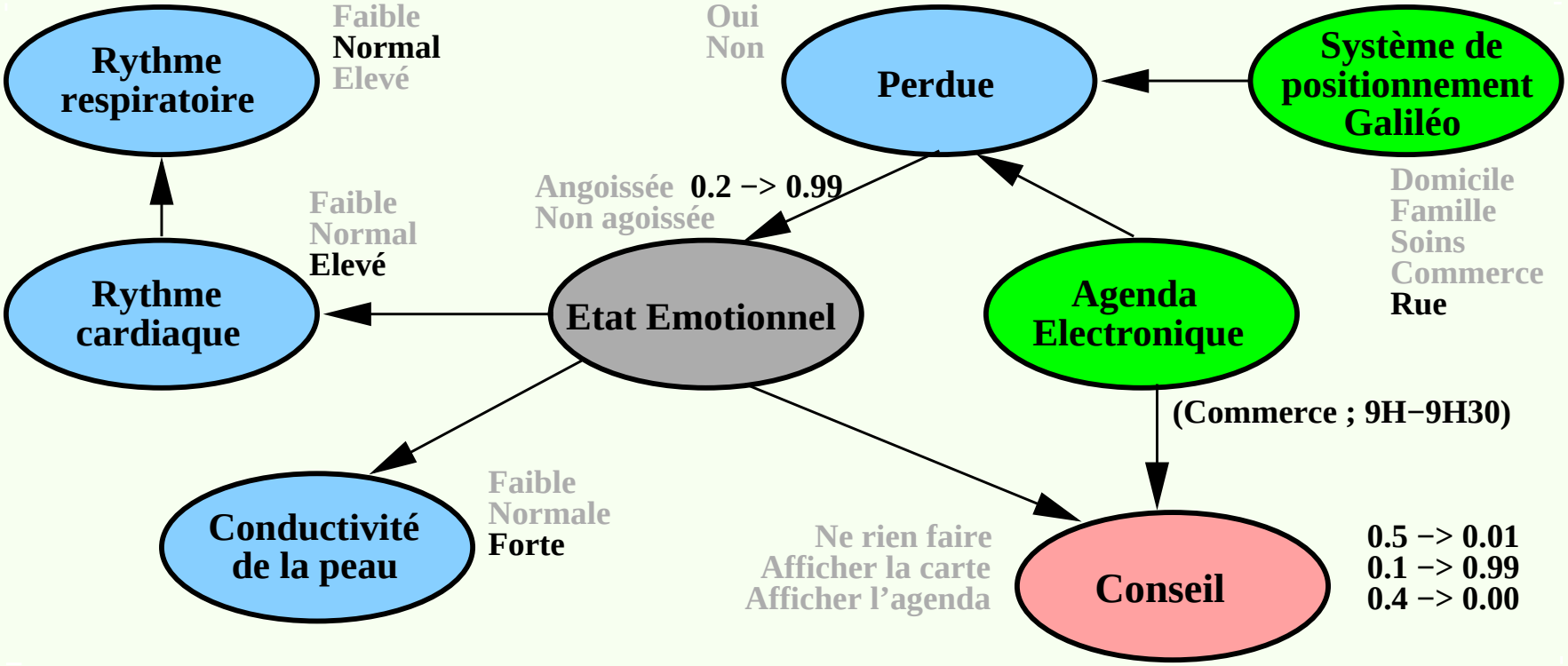
Apprentissage de structure : Pourquoi ?



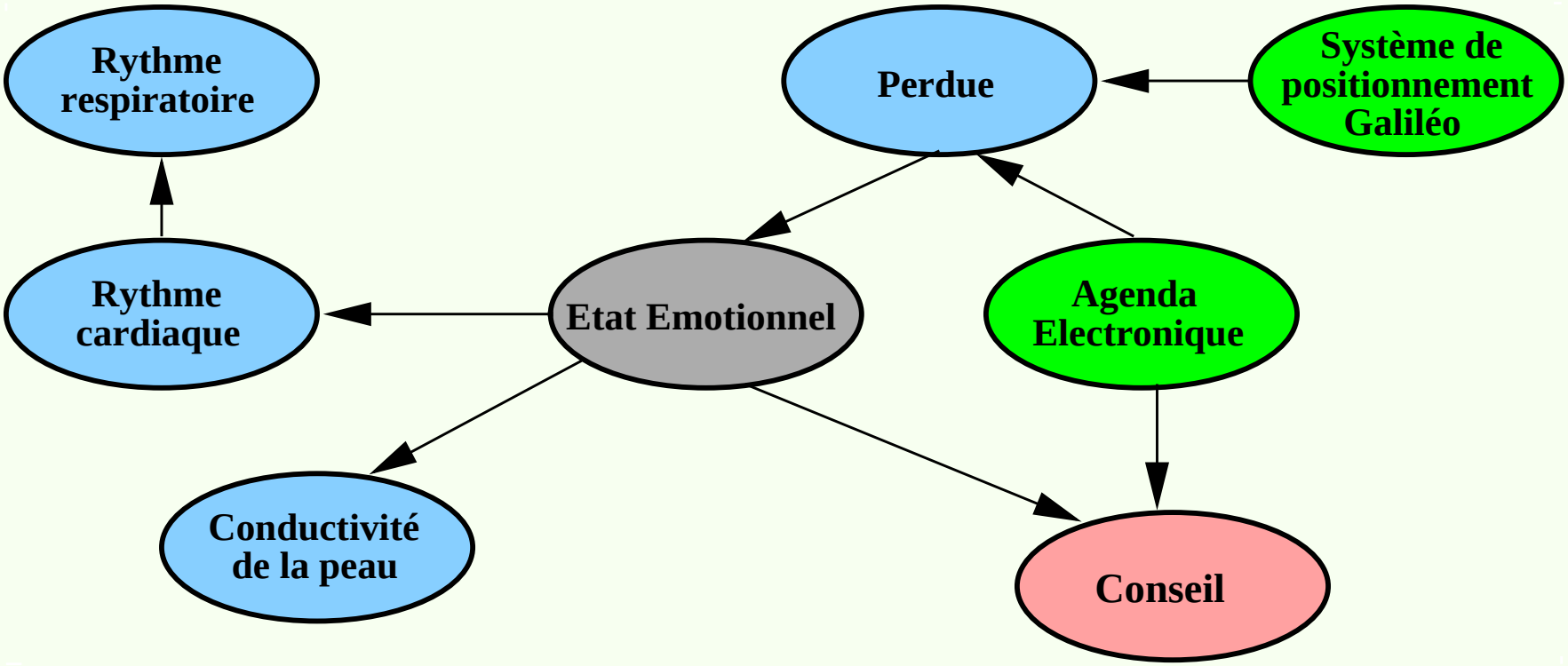
Apprentissage de structure : Pourquoi ?



Apprentissage de structure : Pourquoi ?



Apprentissage de structure : Pourquoi ?



Plan de l'exposé

- Etude empirique d'algorithmes d'apprentissage de structure de Réseaux Bayésiens à partir de **bases** d'exemples **complètement observées**.
  - Utilisation de Benchmarks existants (jouets ou réels)

- Etude empirique d'algorithmes d'apprentissage de structure de Réseaux Bayésiens à partir de **bases** d'exemples **partiellement observées**.
  - Utilisation de Benchmarks existants (jouets ou réels)

Plan de l'exposé

- Etude empirique d'algorithmes d'apprentissage de structure de Réseaux Bayésiens à partir de **bases** d'exemples **complètement observées**.
  - Utilisation de Benchmarks existants (jouets ou réels)
- Etude empirique d'algorithmes d'apprentissage de structure de Réseaux Bayésiens à partir de **bases** d'exemples **partiellement observées**.
  - Utilisation de Benchmarks existants (jouets ou réels)
- **Génération de bases d'exemples incomplètes**

Plan de l'exposé

1) Apprentissage de structure de RB avec D. complètes
   - Etat de l'art
   - Notre proposition
   - Résultats
2) Apprentissage de RB avec D. incomplètes
   - Etat de l'art
   - Nos propositions : MWST-EM et SEM+T
   - Résultats et interprétations
   - Application à la classification : TAN-EM
3) Conclusion et Perspectives

Apprentissage de structure

Deux classes d'algorithmes :

Méthodes à base de contraintes

Utiliser des **tests statistiques** pour identifier les **indépendances conditionnelles** des attributs et en déduire une structure.

- → test du χ^2
- → test du rapport de vraisemblance
- → information mutuelle
- → étude des corrélations...

Méthodes à base de score

Maximiser une mesure/approximation de **la vraisemblance** marginale dans un espace de structures.

- → scores *BD*, *BDe*, *BDeu*
- → critères *AIC* et *BIC*
- → *M*inimum *D*escription *L*ength
- → approximation de Laplace au *n*-ième ordre
- → *AICc*, *ICL* ...

Espaces de recherche

Recherche d'un bon réseau bayésien :

- **exhaustive = impossible** / taille de l'espace

$$n_{\mathrm{DAG}}(n) = n^{2^{\mathcal{O}(n)}}$$

$n_{\mathrm{DAG}}(5) = 29\,281$, $n_{\mathrm{DAG}}(8) = 783\,702\,329\,343$, $n_{\mathrm{DAG}}(10) \simeq 4,175 \times 10^{18}$

Espaces de recherche

Recherche d'un bon réseau bayésien :

- **exhaustive = impossible** / taille de l'espace

$$n_{\mathrm{DAG}}(n) = n^{2^{\mathcal{O}(n)}}$$

$n_{\mathrm{DAG}}(5) = 29\,281$, $n_{\mathrm{DAG}}(8) = 783\,702\,329\,343$, $n_{\mathrm{DAG}}(10) \simeq 4,175 \times 10^{18}$

Dans quel espace ?

- espace des **arbres**
- espace des DAG respectant un **ordre d'énumération**
- espace des DAG avec un **nombre max. de parents**

- espace des **DAG** (heuristiques de parcours)
- espace des **classes d'équivalence de Markov** (heur.)

→ **CPDAG** : représentants 'minimaux' des classes d'équivalence

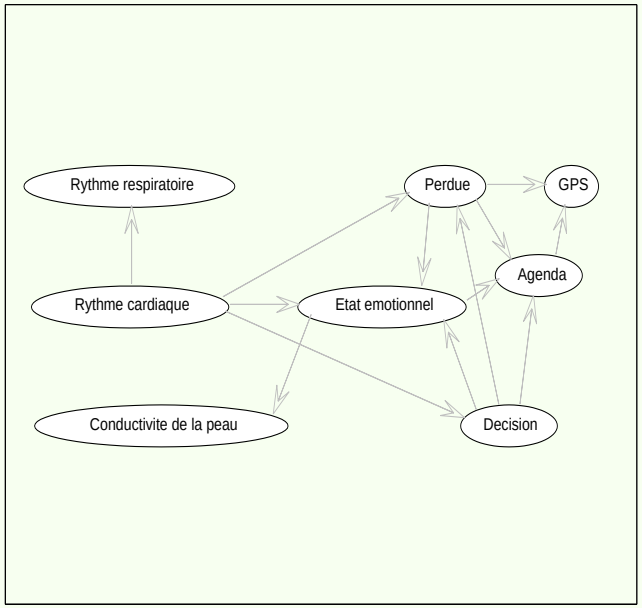
Apprentissage de structure à partir de données complètes

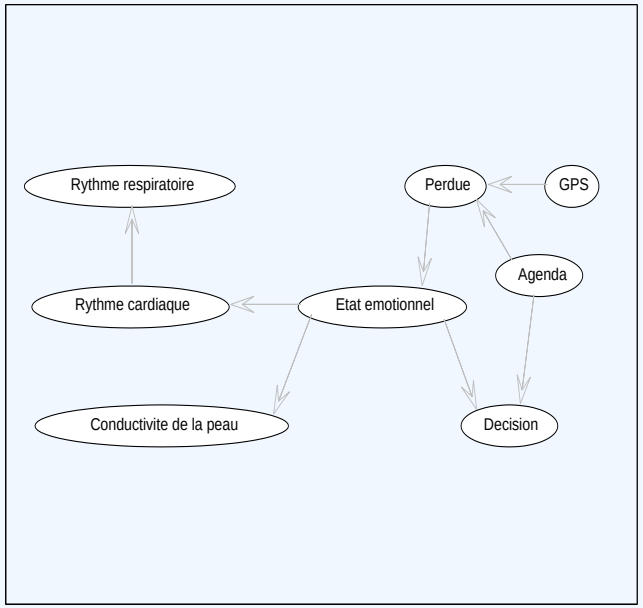
	Méthodes testées	Espaces	An
Contraintes	**PC**	CPDAG	93
	BNPC	CPDAG	02
Scores	**MWST**	arbres	94
	K2	ordre fixe	92
	GS	DAG	95
	GES	CPDAG	02

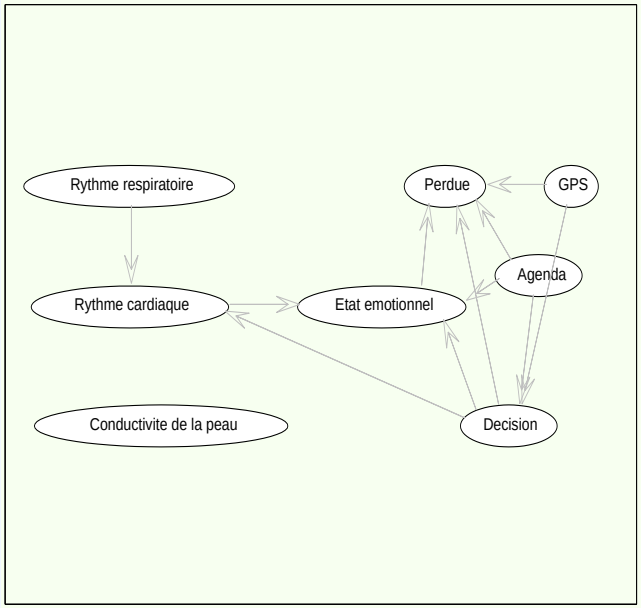
Apprentissage de structure à partir de données complètes

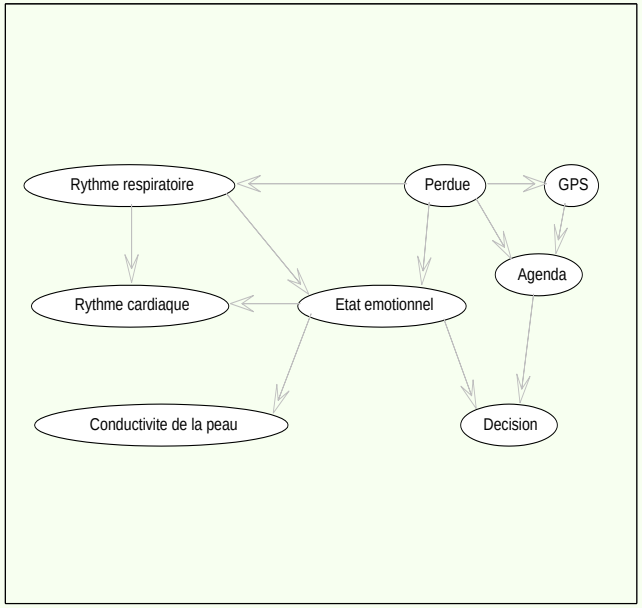
	Méthodes testées	Espaces	An	Autres méthodes	An
Contraintes	**PC**	CPDAG	93	IC/IC*	93,00
				FCI/QFCI/AFCI	00,03,06
	BNPC	CPDAG	02	RAI	05
				MFC	06
Scores	**MWST**	arbres	94		
	K2	ordre fixe	92	MCMC	00
				Fourmis	02
	GS	DAG	95	Algos G	04,06
	GES	CPDAG	02	K2+GSno+GES	06
Cont.		CPDAG		Grow-Shrink	03
				MaxMin HillC	05

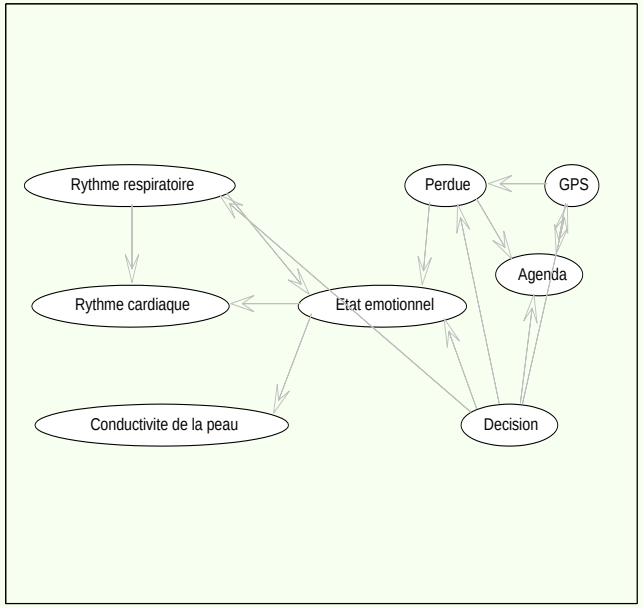
Problèmes d'initialisation : K2

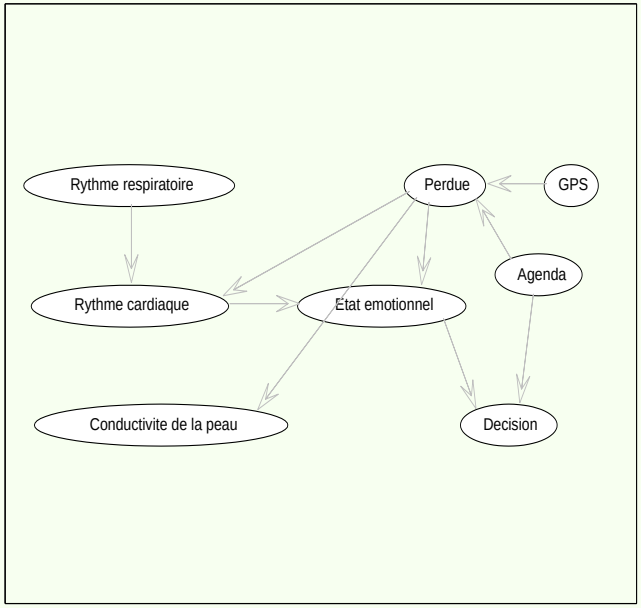






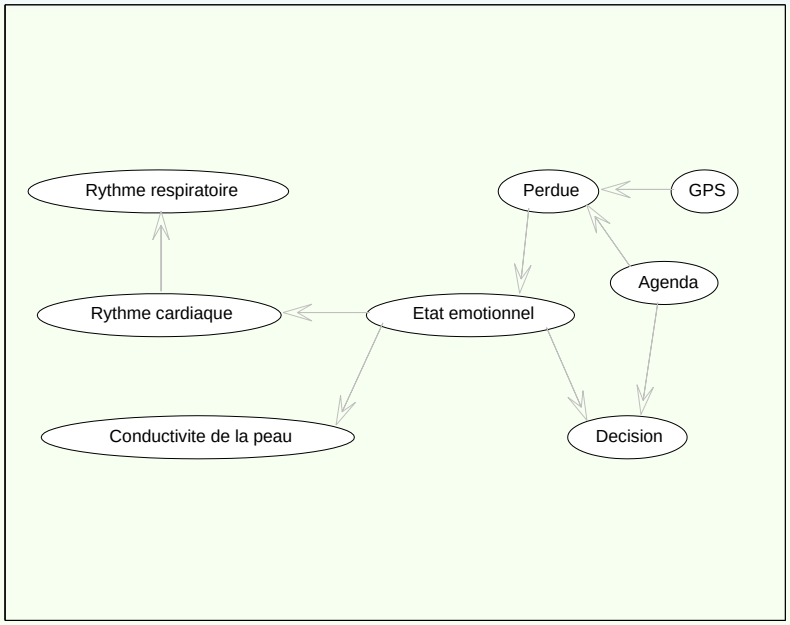


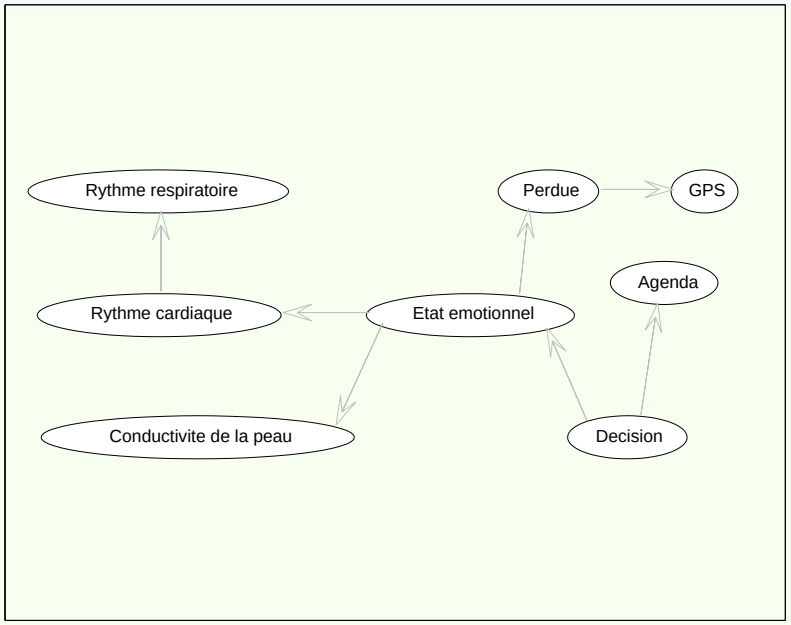




Problèmes d'initialisation : K2

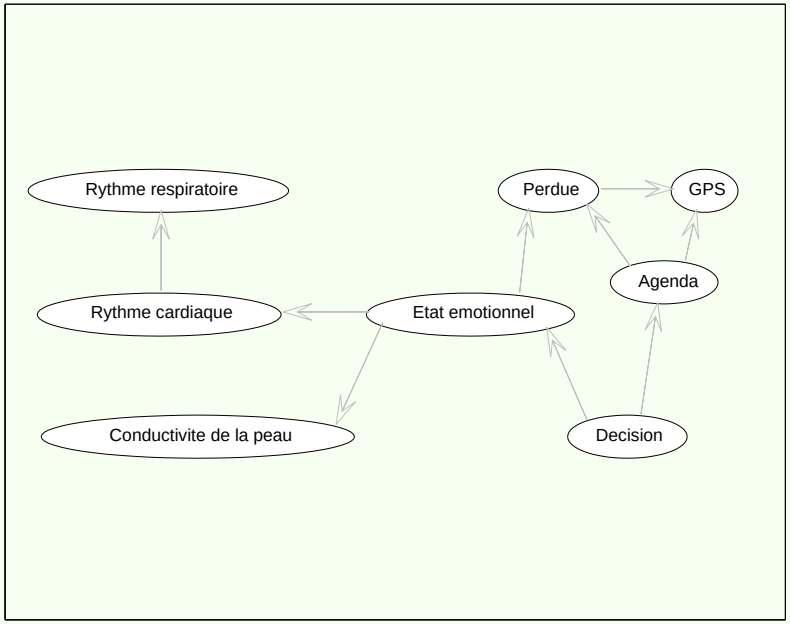
Orig.

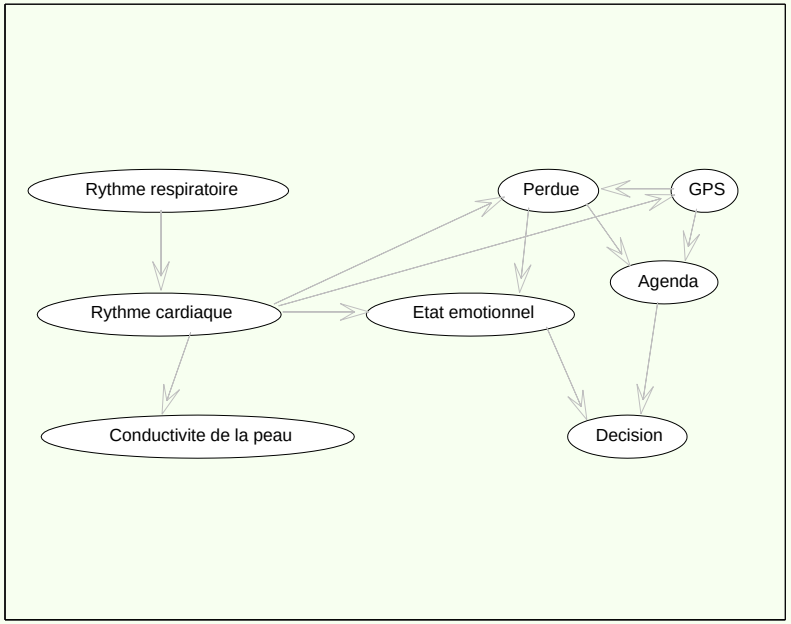




MWST

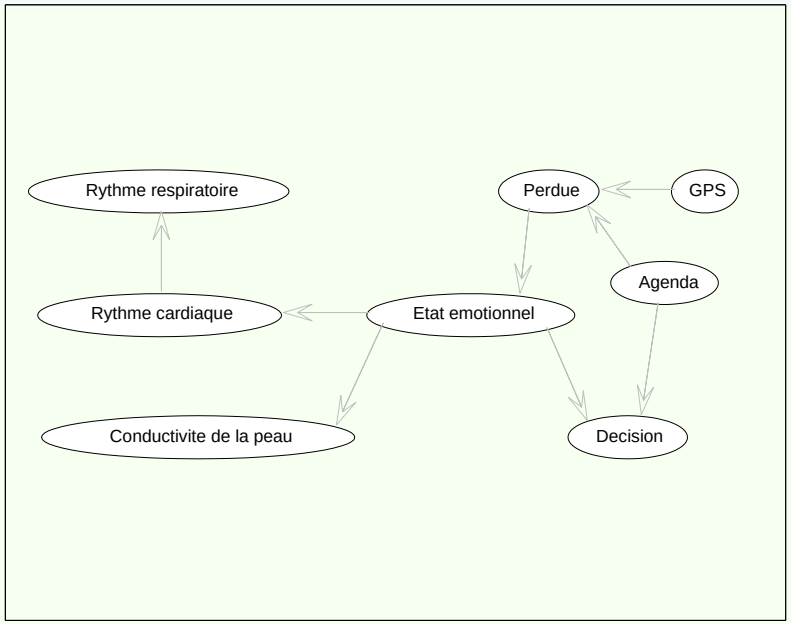
K2+T



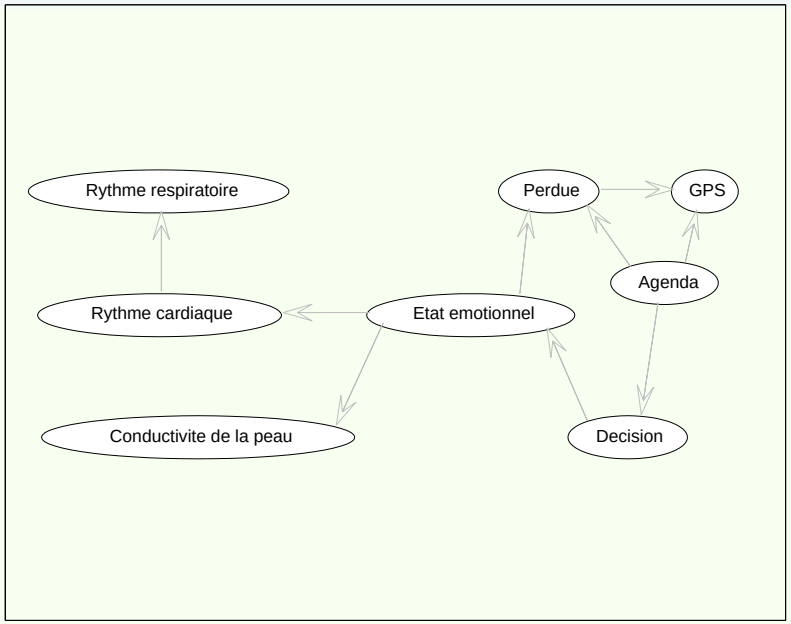


K2-T

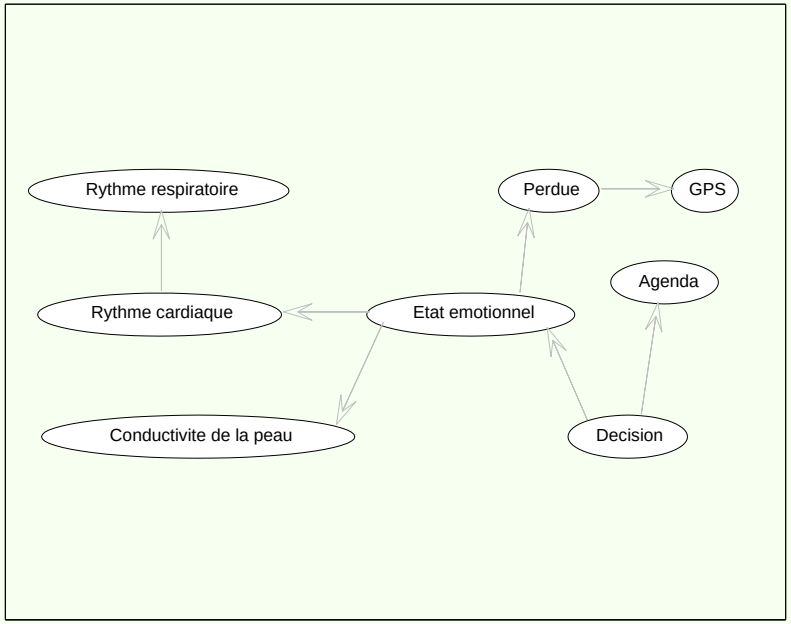
Problèmes d'initialisation : Greedy Search



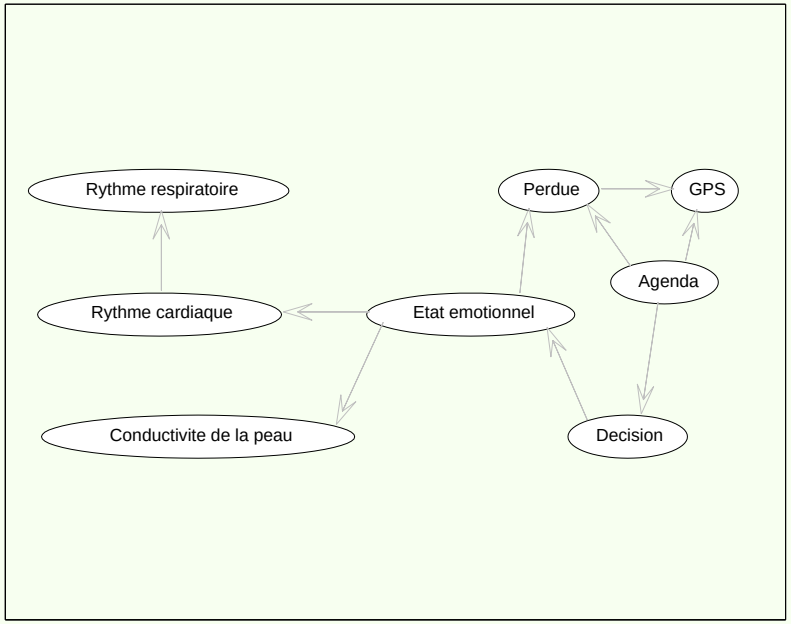
Orig.



GS



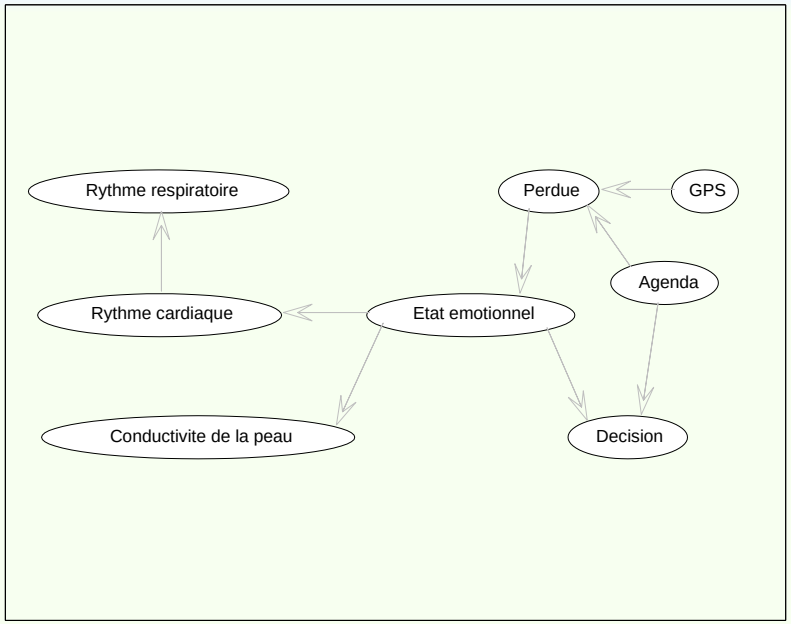
MWST

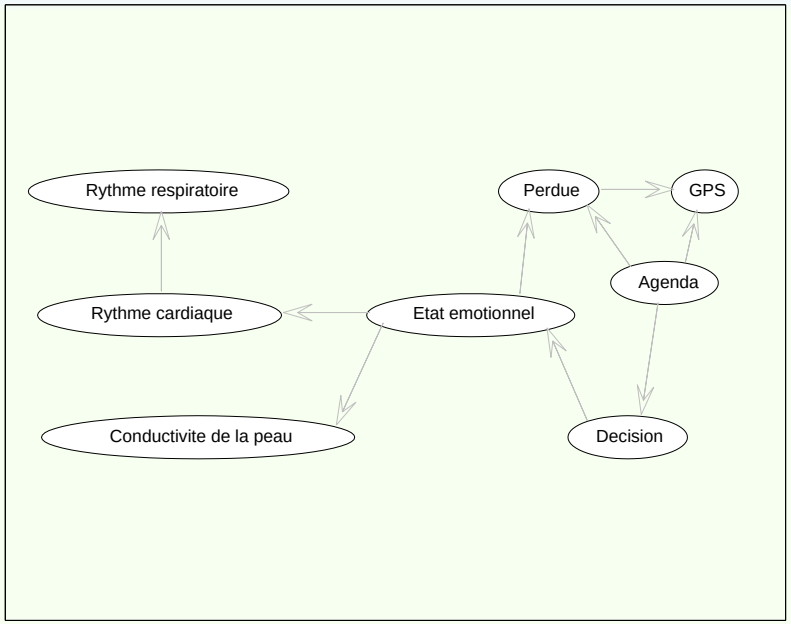


GS+T

Problèmes d'initialisation : Greedy Search

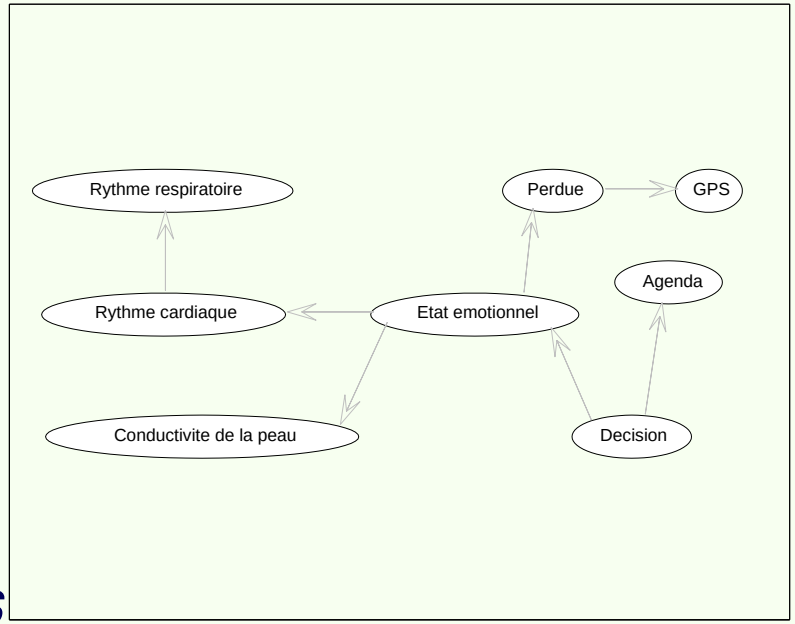
Orig.

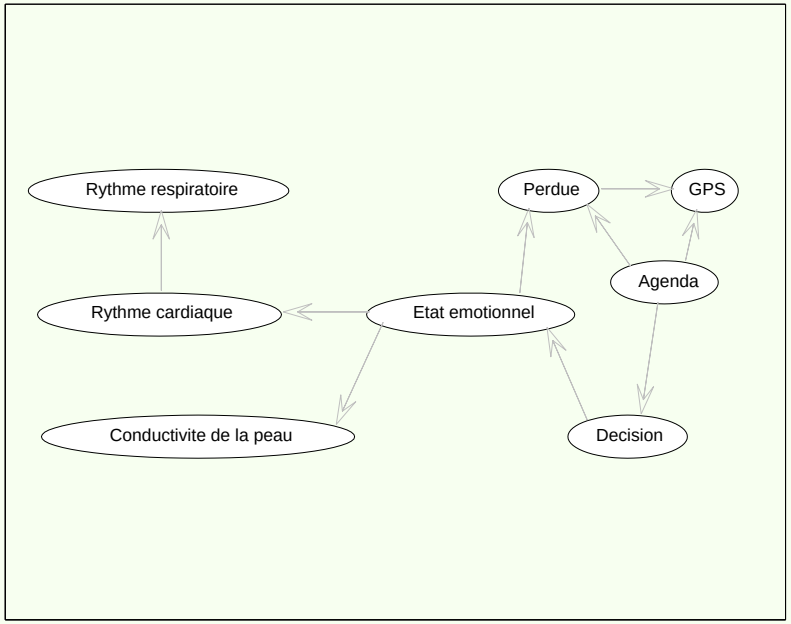




$6s$

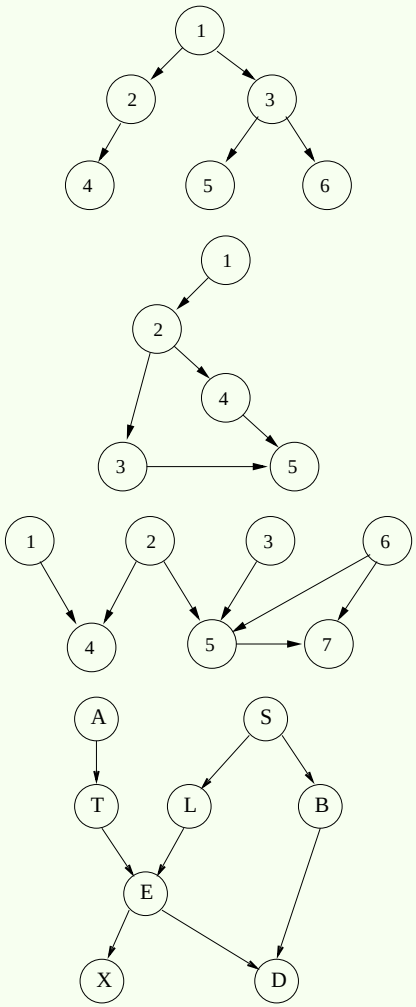
$0,2s$

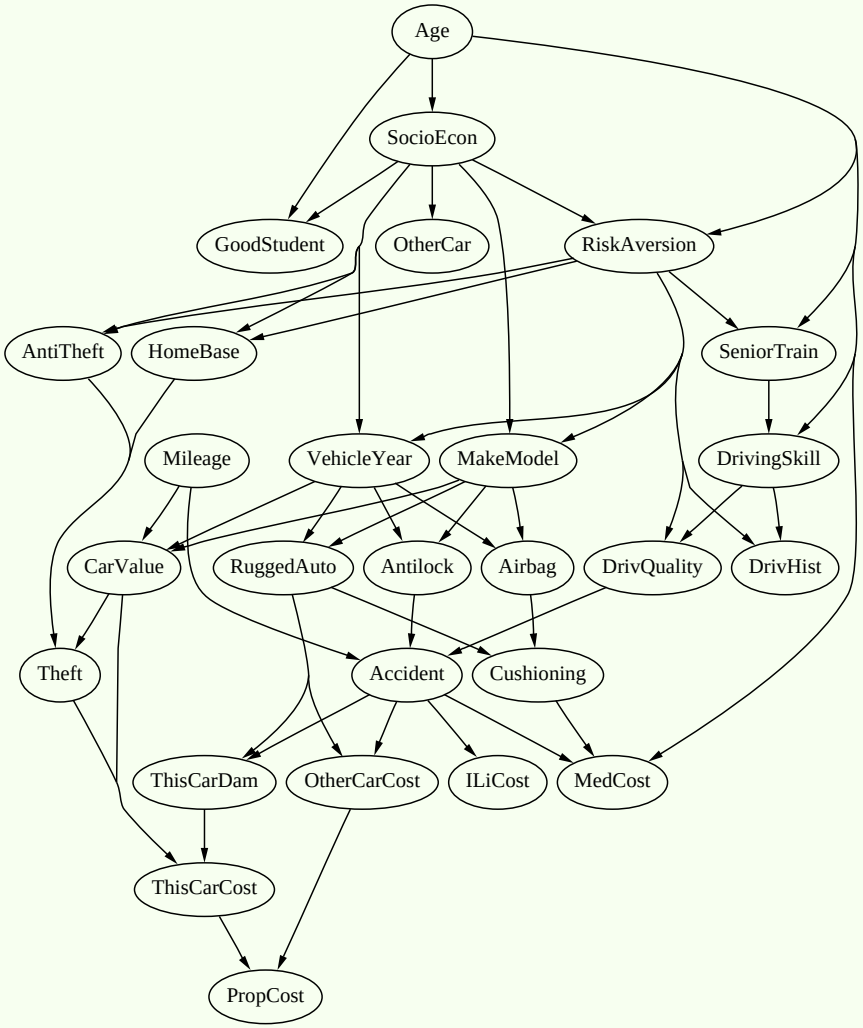




$2,5s$

Des réseaux et des tests...





→ retrouver ces structures à partir de bases de **différentes tailles**.
314 problèmes synthétiques et 17 problèmes réels (UCI, Statlog).

Des réseaux et des tests...

Nous avons testé :

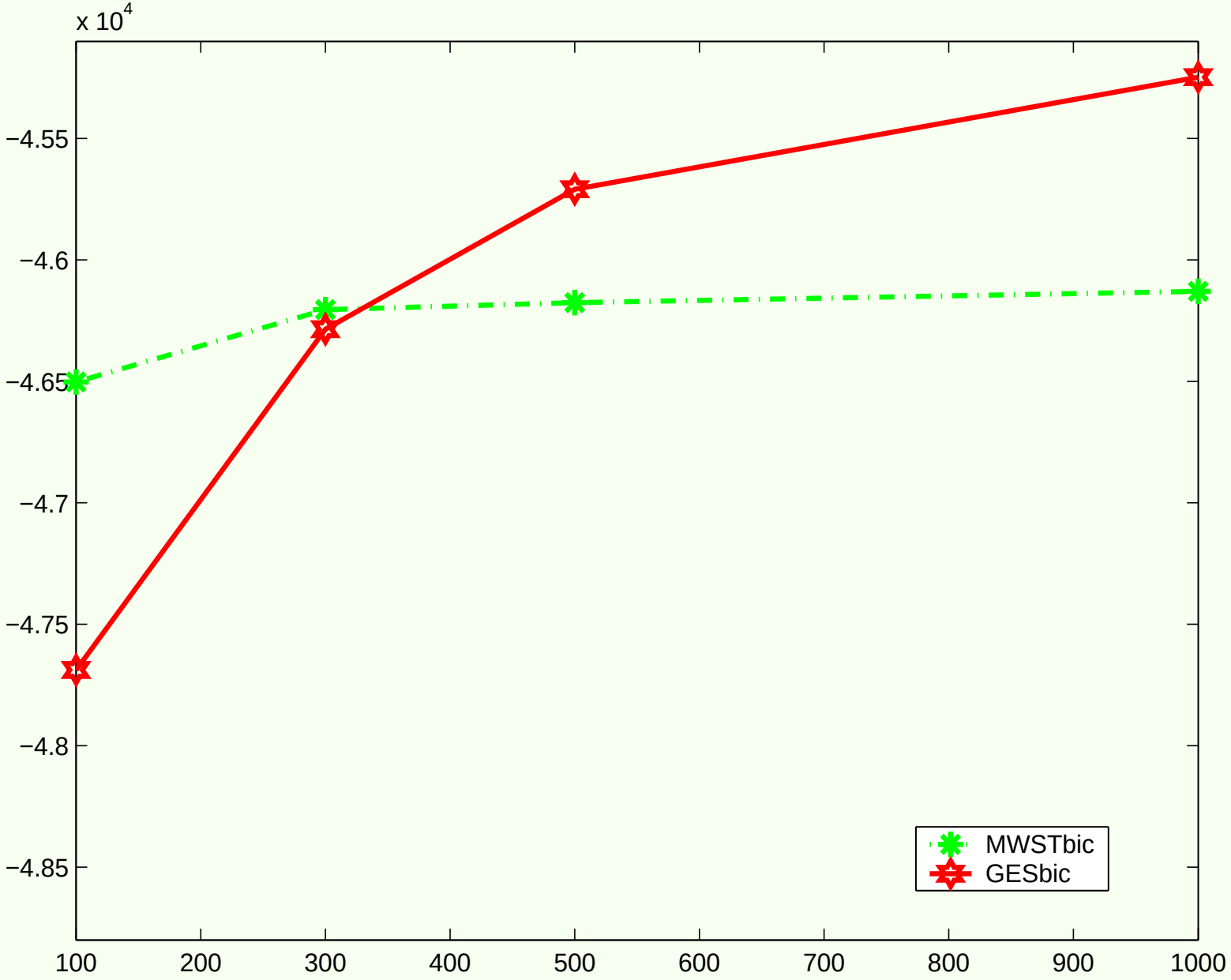
- performances BIC
- performances KL
- performances distance d'édition
- influence de la taille de la base d'exemples
- recherche de structures complexes
- efficacité à retrouver une dépendance faible
- stabilité BIC
- stabilité KL
- temps de calcul
- utilisation des RB en classification

Des réseaux et des tests...

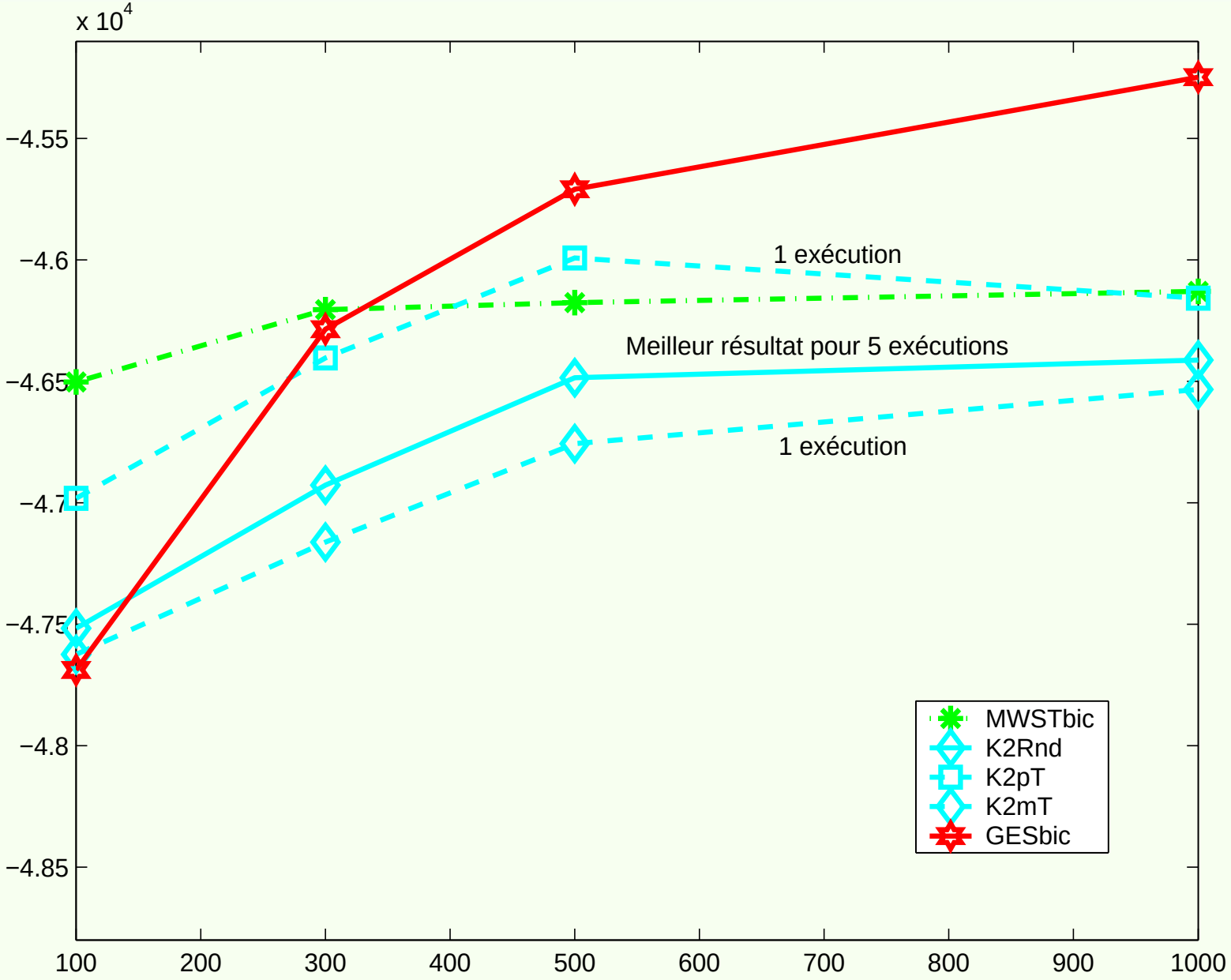
Nous avons testé :

- **performances BIC**
- performances KL
- performances distance d'édition
- **influence de la taille de la base d'exemples**
- recherche de structures complexes
- efficacité à retrouver une dépendance faible
- stabilité BIC
- stabilité KL
- temps de calcul
- **utilisation des RB en classification**

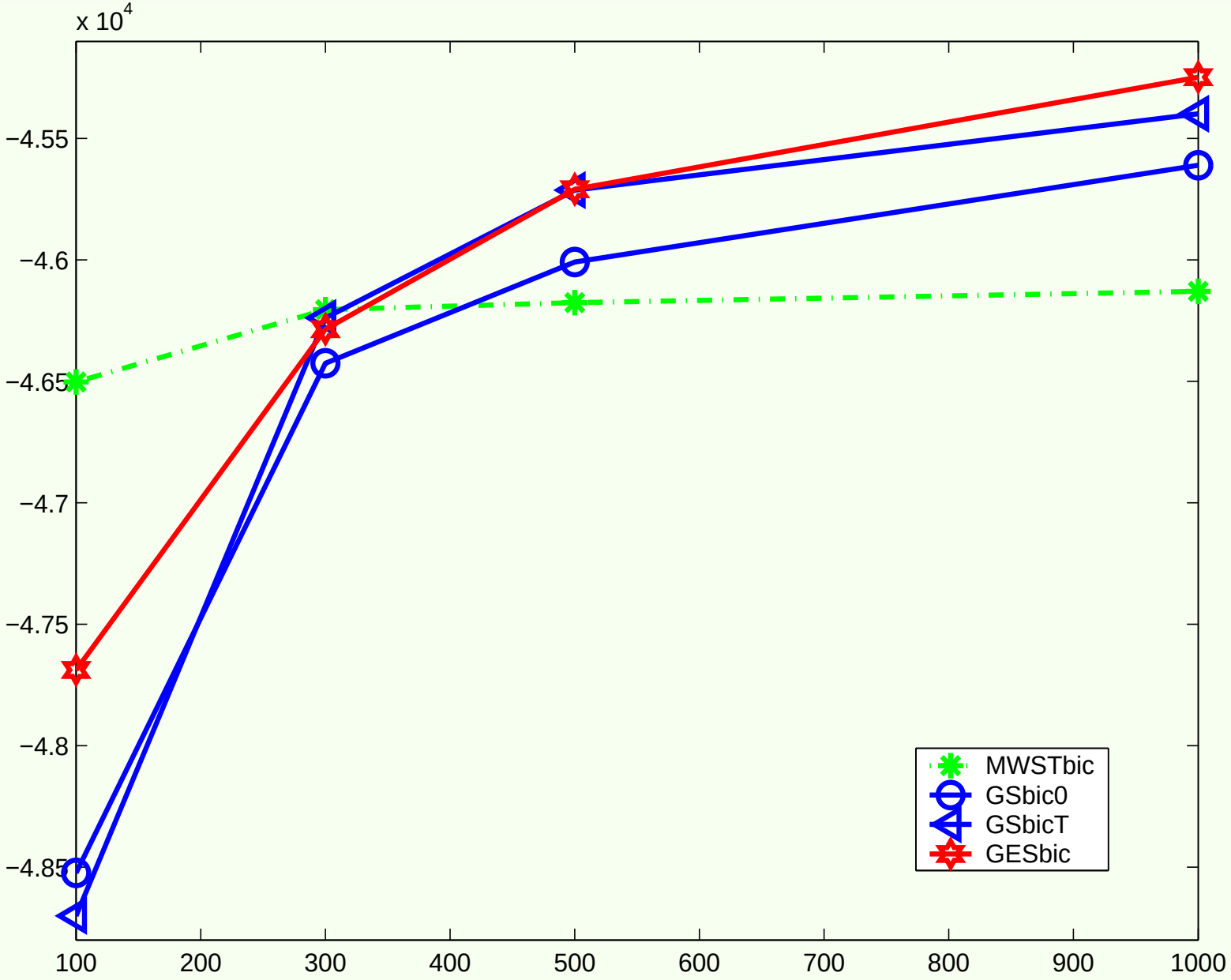
Performances BIC



Performances BIC

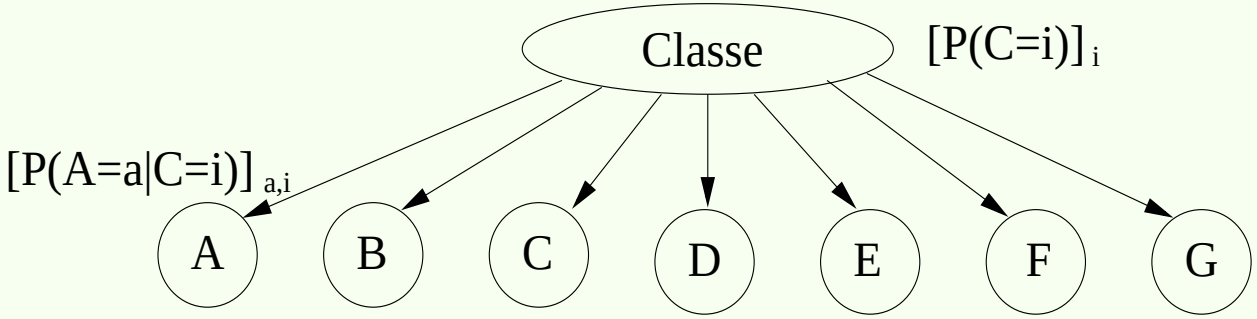


Performances BIC



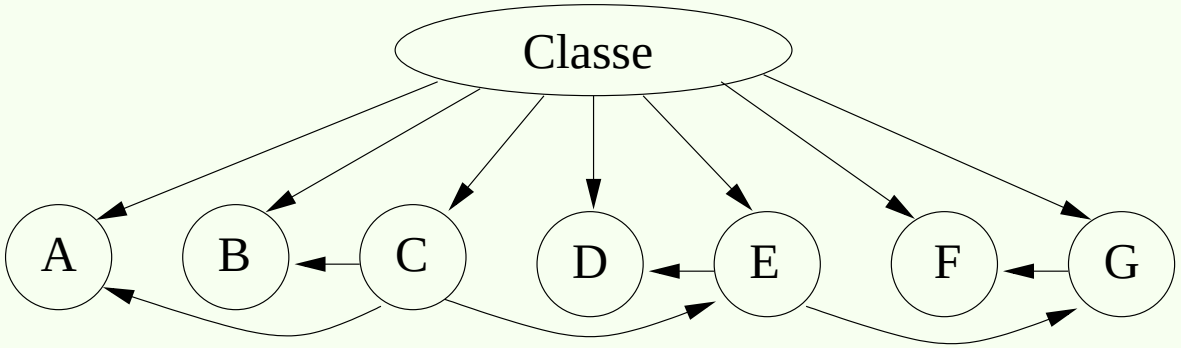
Le réseau Naïf

Supposons que les *attributs soient indépendants 2 à 2 conditionnellement à la classe*, alors

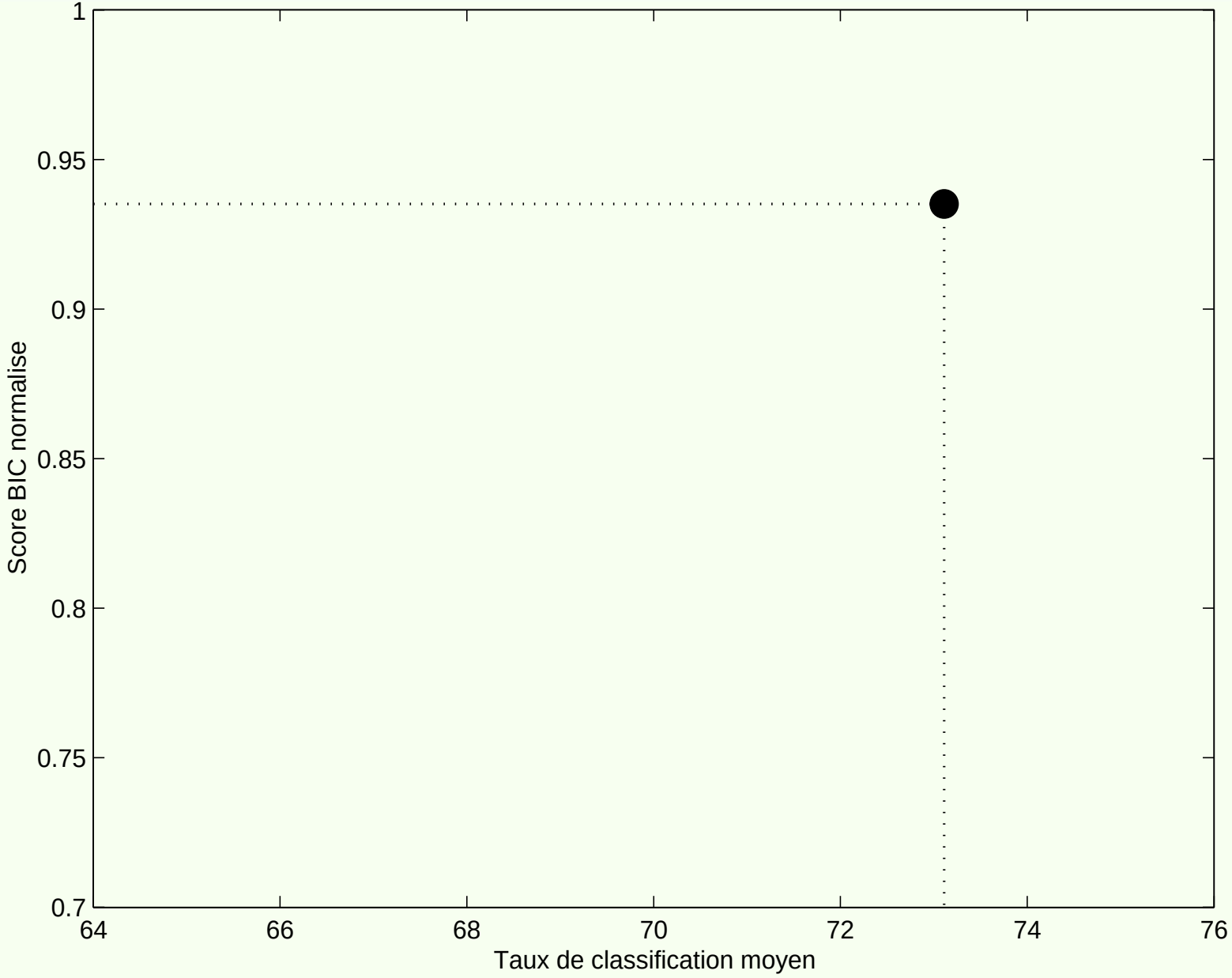


Le réseau Naïf augmenté par un arbre (TAN)

Supposons que les *attributs* NE *soient* PLUS *indépendants 2 à 2 conditionnellement à la classe*, alors

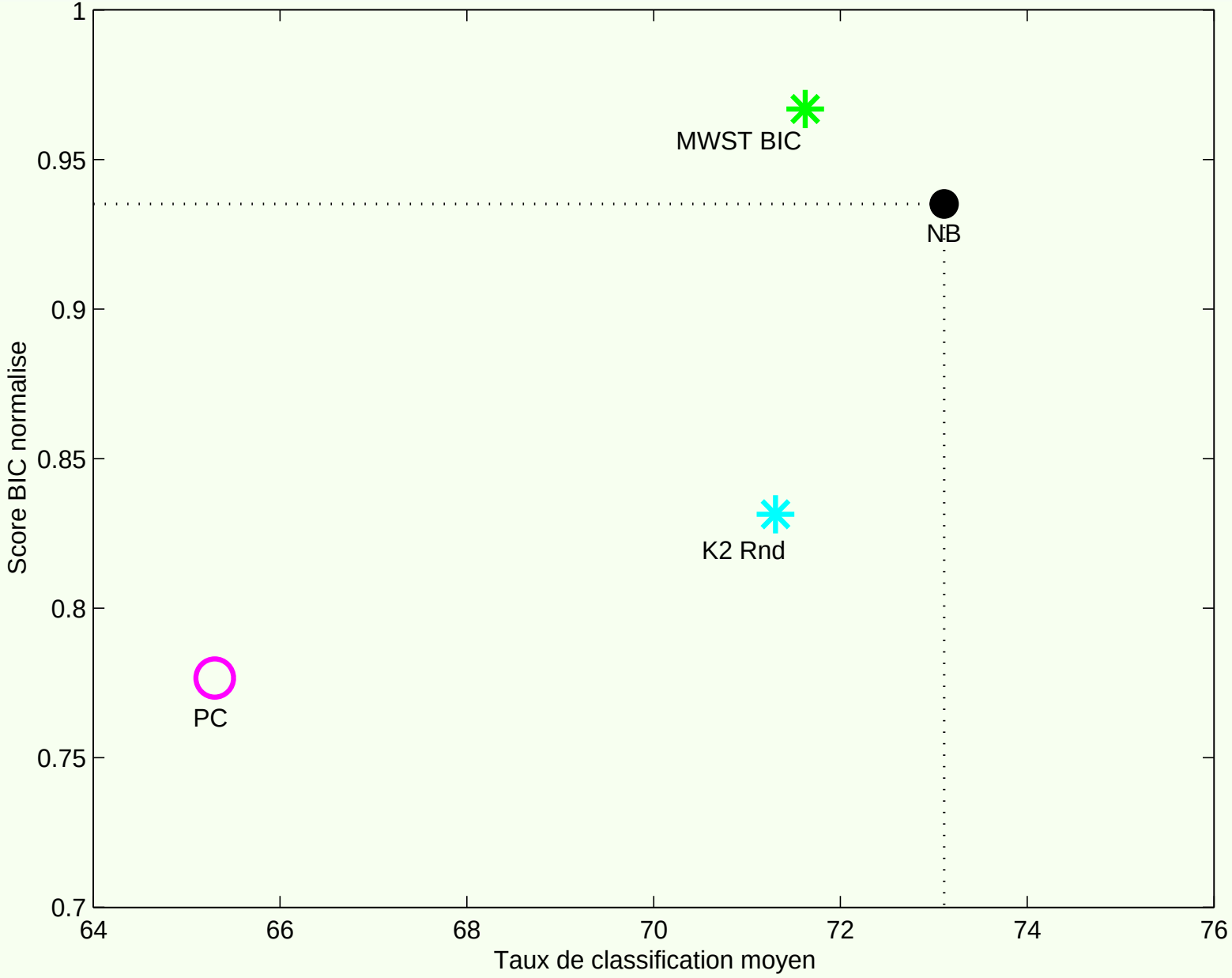


Performance en classification



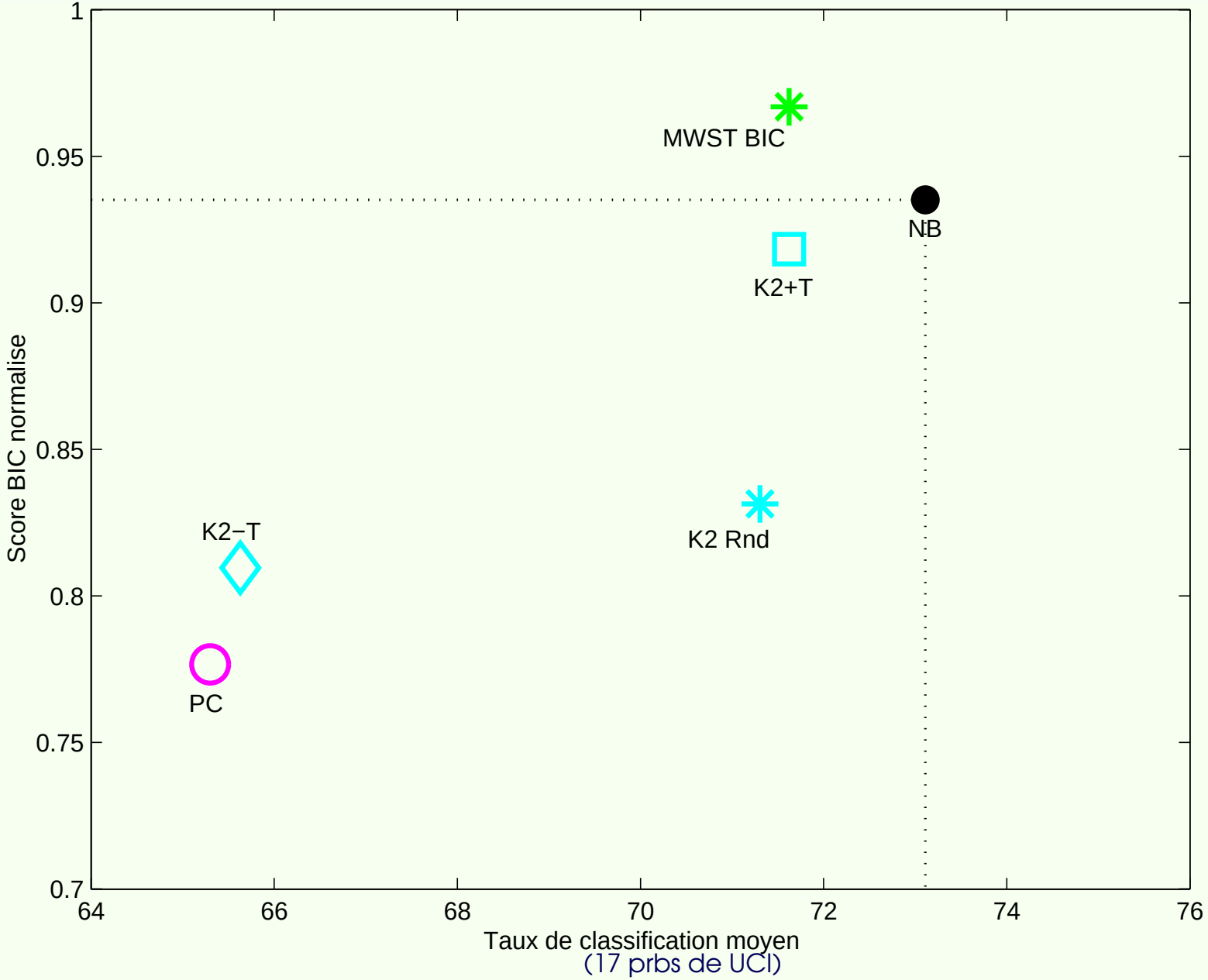
(17 prbs de UCI)

Performance en classification

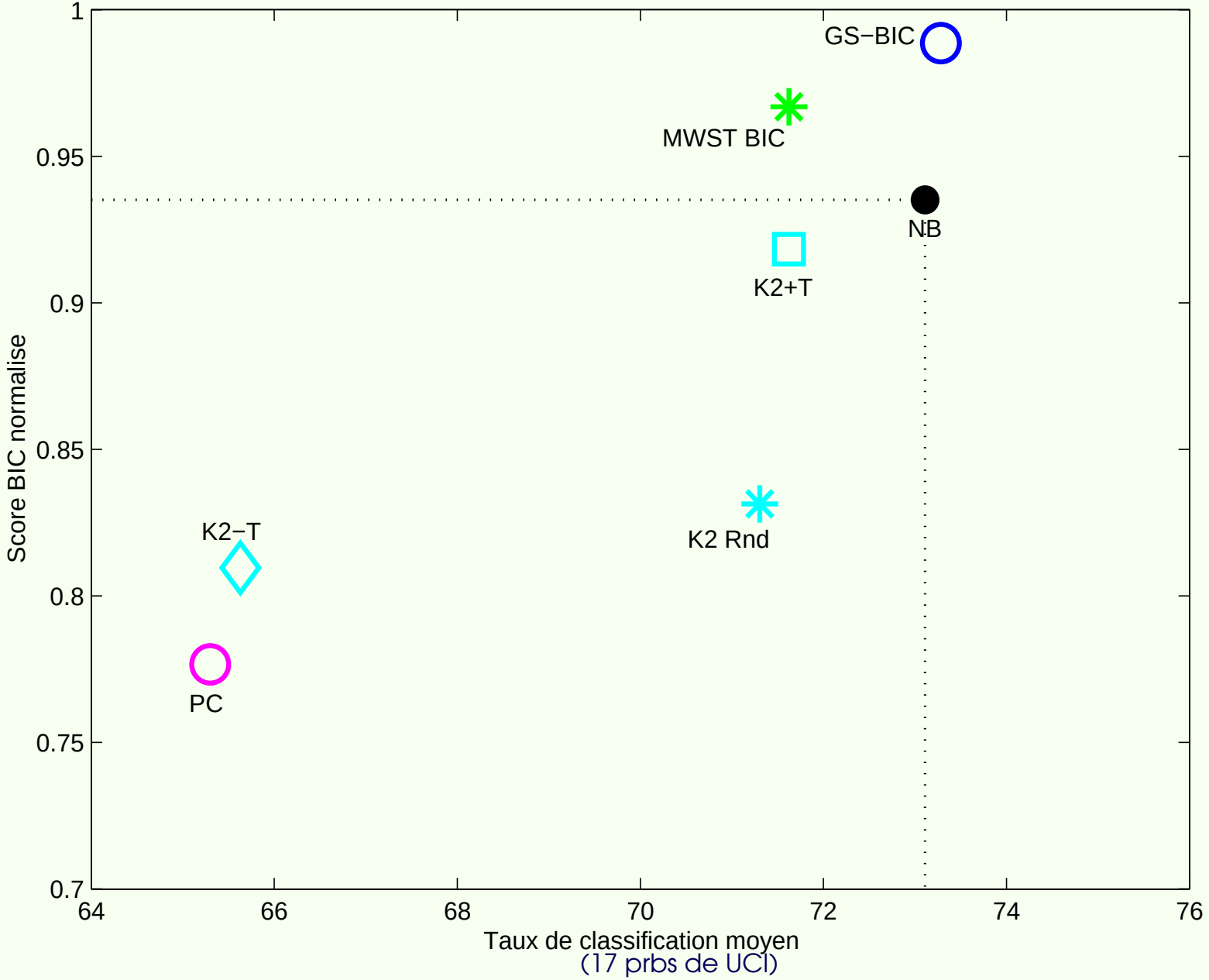


(17 prbs de UCI)

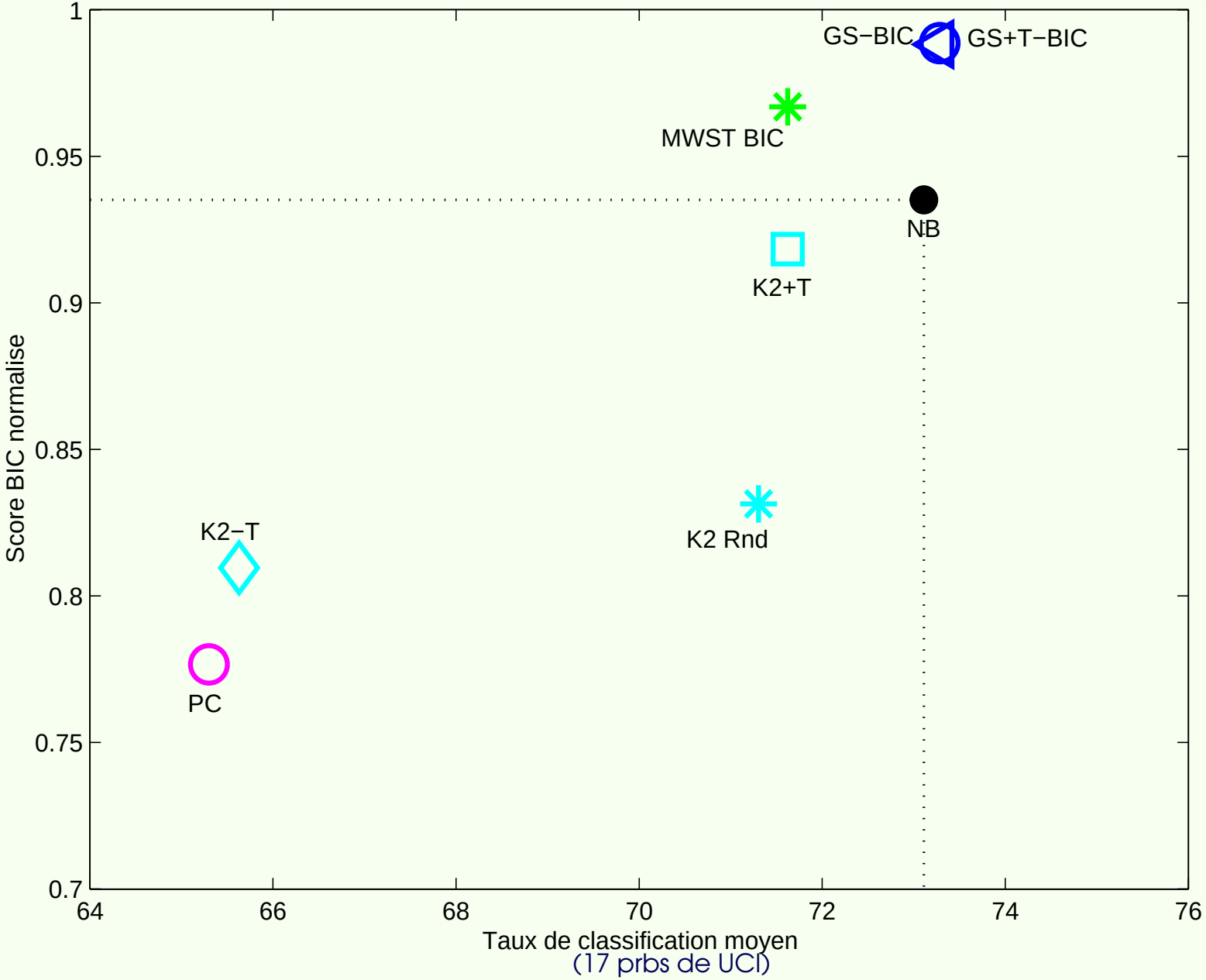
Performance en classification



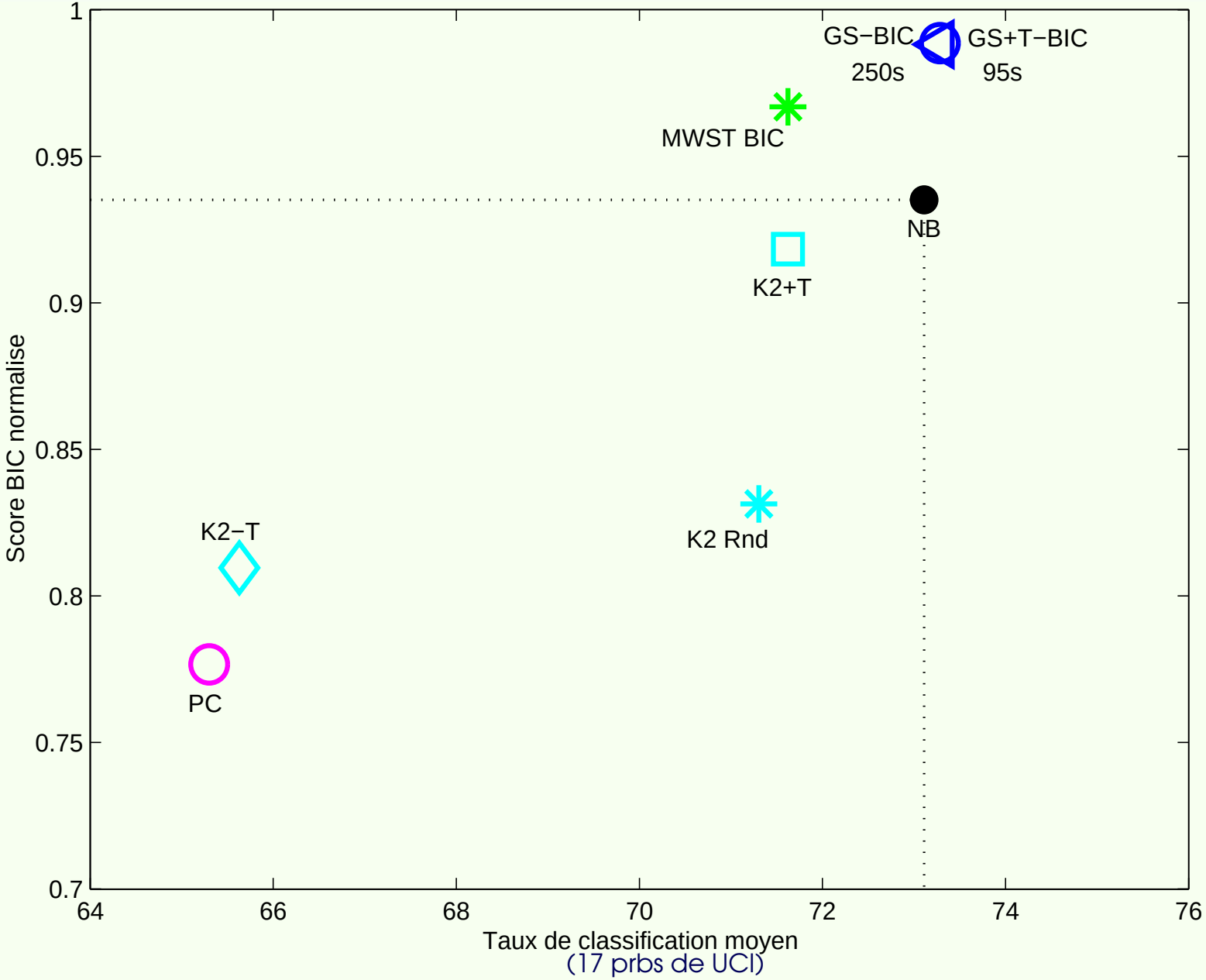
Performance en classification



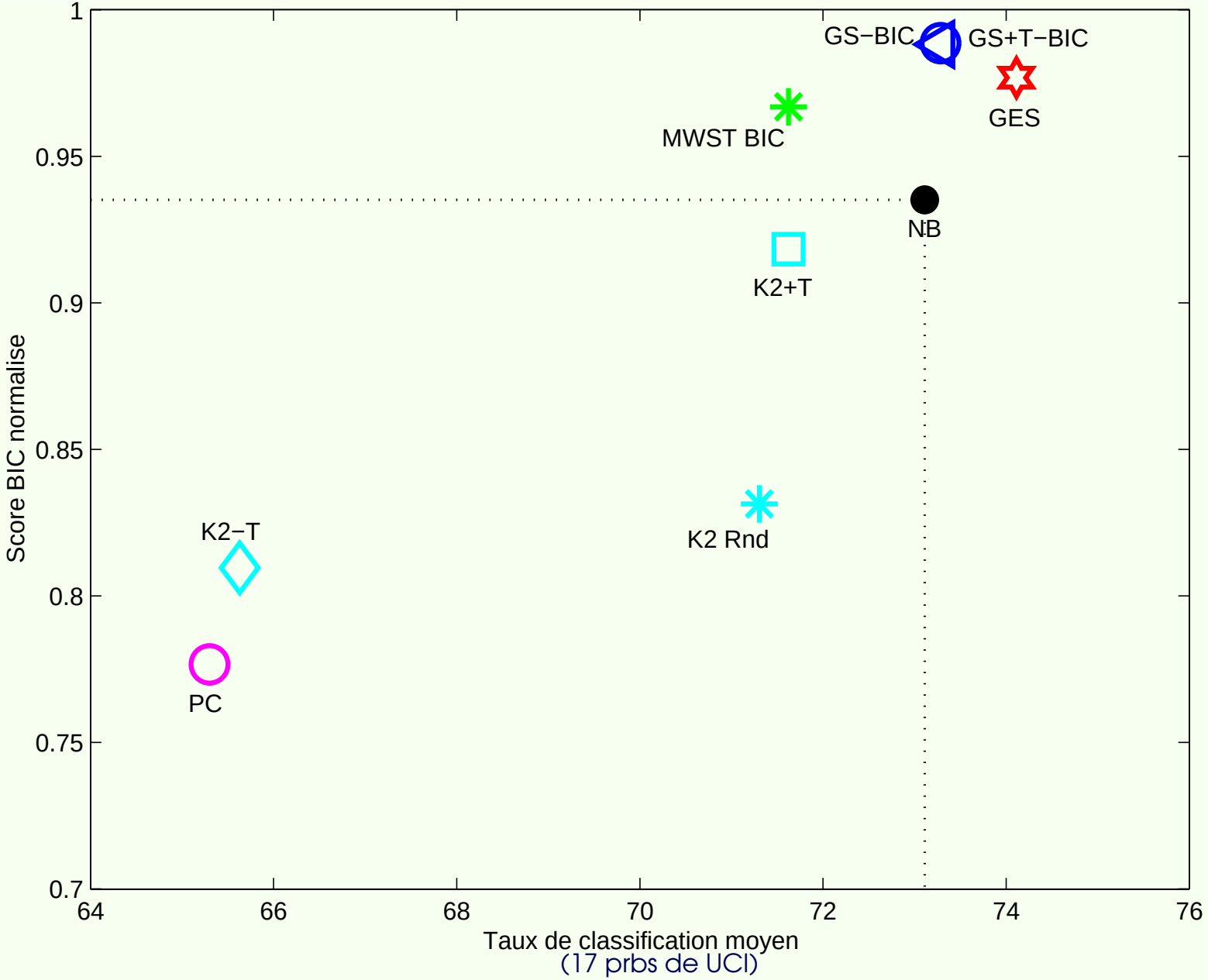
Performance en classification



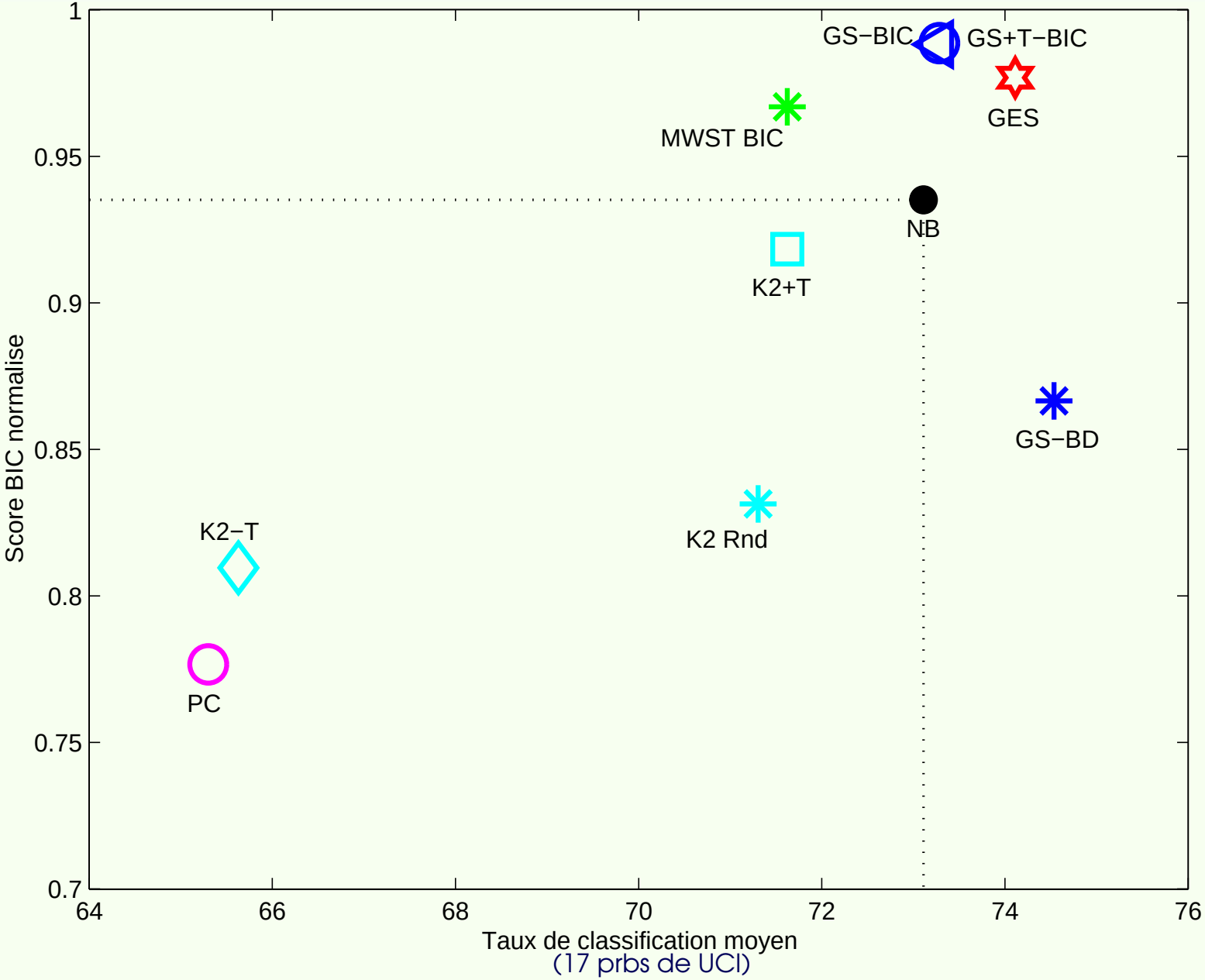
Performance en classification



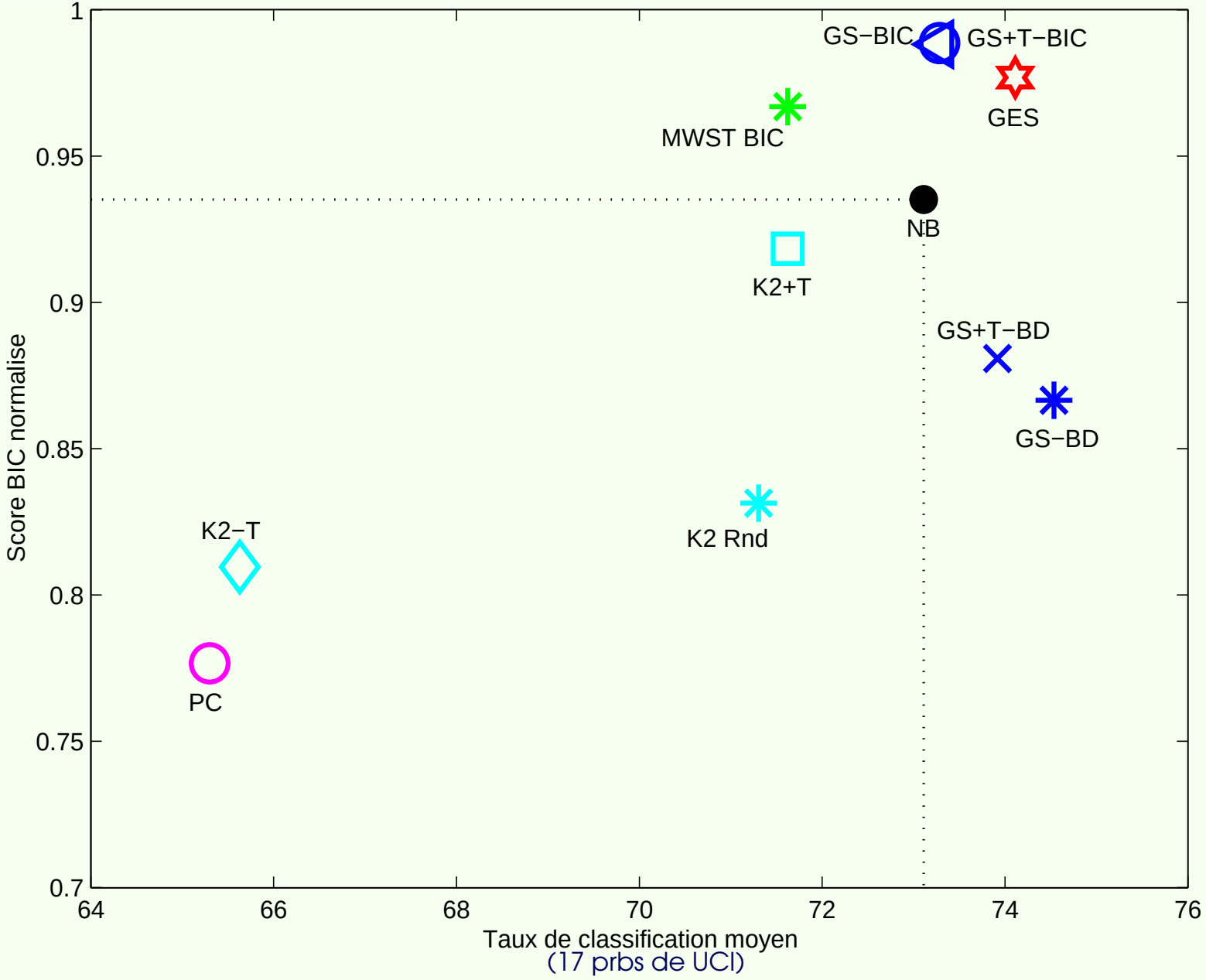
Performance en classification



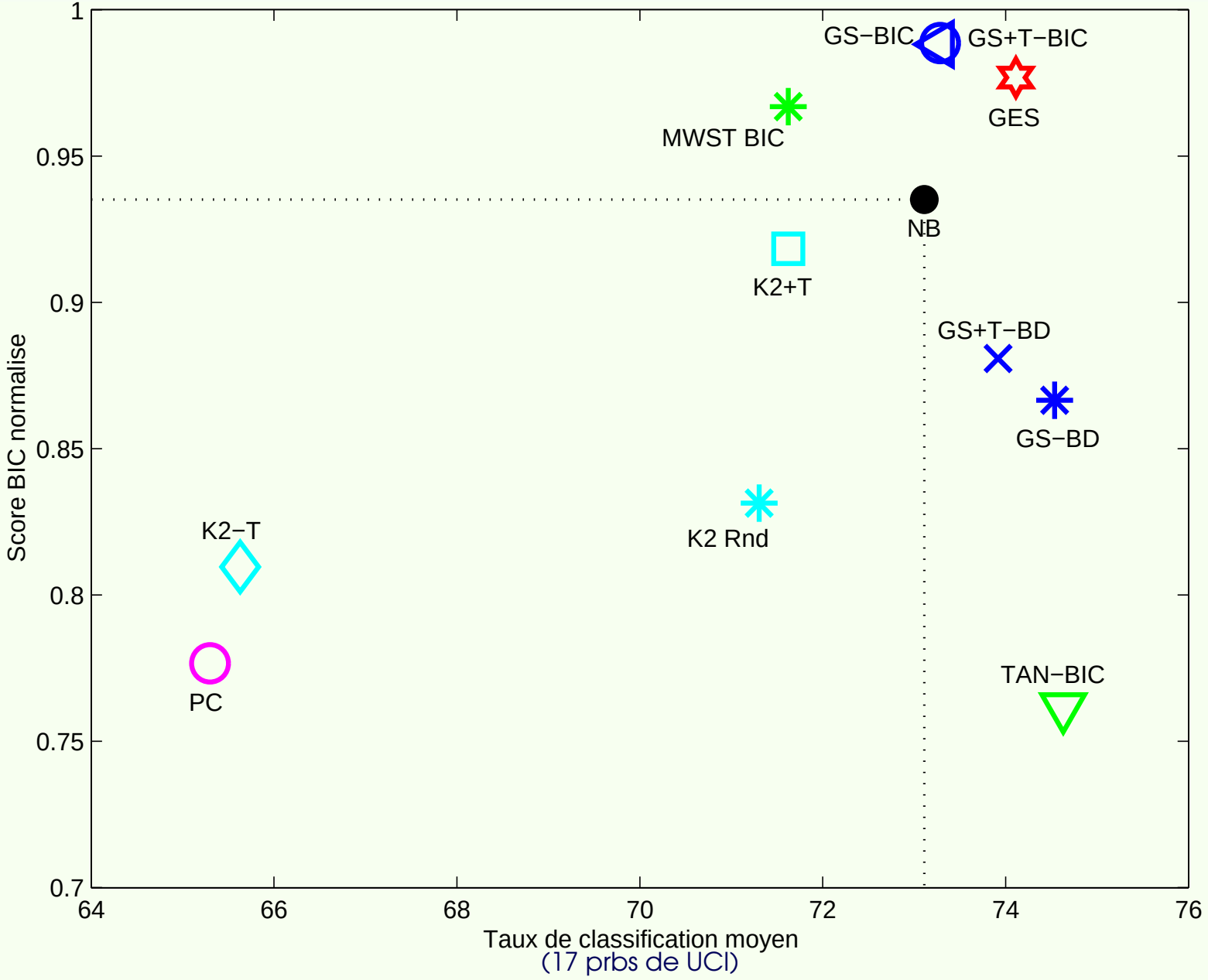
Performance en classification



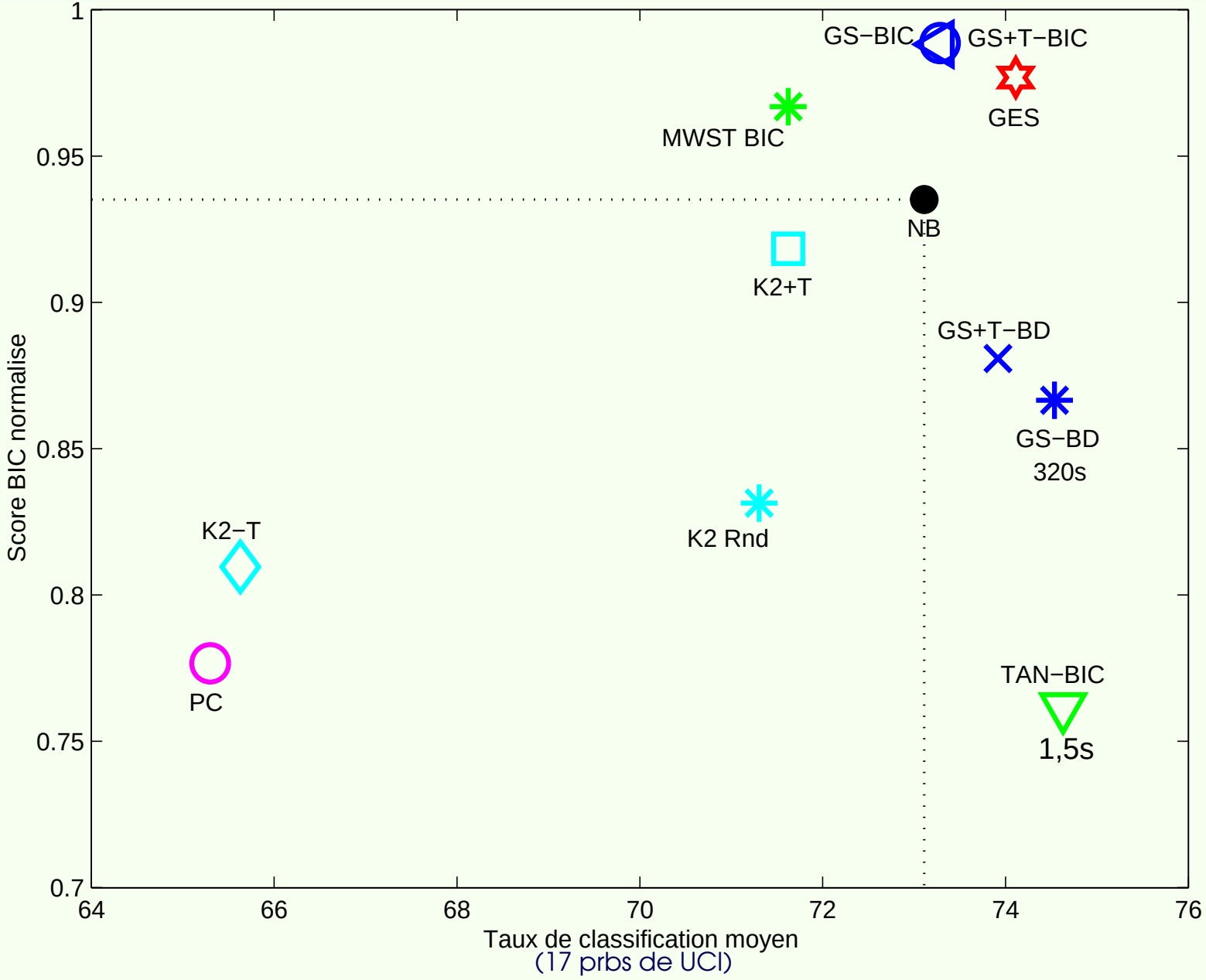
Performance en classification



Performance en classification



Performance en classification



Plan de l'exposé

1) Apprentissage de structure de RB avec D. complètes
   - Etat de l'art
   - Notre proposition
   - Résultats
2) Apprentissage de RB avec D. incomplètes
   - Etat de l'art
   - Nos propositions : MWST-EM et SEM+T
   - Résultats et interprétations
   - Application à la classification : TAN-EM
3) Conclusion et Perspectives

Réseaux Bayésiens et données incomplètes

- Les RB peuvent naturellement effectuer de l'inférence avec des données manquantes.
- Des heuristiques (*EM, MCMC...*) ont été adaptées pour l'estimation des paramètres.

Qu'en est-il pour l'apprentissage de structure ?

Réseaux Bayésiens et données incomplètes

- Les RB peuvent naturellement effectuer de l'inférence avec des données manquantes.
- Des heuristiques (*EM, MCMC...*) ont été adaptées pour l'estimation des paramètres.

Qu'en est-il pour l'apprentissage de structure ?

- AMS-EM : recherche gloutonne de DAG (Friedman'97) **SEM**
- BS-EM : recherche gloutonne de DAG (Friedman'98)
- Algo. evolutionnaires et *MCMC* (Myers'99)
- Hybrid Independence Test (Dash'03)

Motivation

Vu précédement (RFIA'04)

- MWST : bon rapport complexité/performance
- GS+T : permet de gagner en stabilité
- SEM = GS avec BIC/MDL + EM

MWST-EM = MWST + EM

- Le rapport complexité/perfs. de MWST-EM est-il bon ?
- MWST-EM est-t'il une bonne initialisation pour SEM ?
  $\longrightarrow$ **SEM+T**

Donner un score à partir d'une base incomplète

Soit $\mathcal{S}(\mathcal{M}|\mathbf{D}_C)$, un critère de score pour un modèle $\mathcal{M}$ et des données complètes $\mathbf{D}_C$

Comment évaluer $\mathcal{S}$ lorsque la base est incomplète ?

$$\mathbf{D} = < \mathbf{O}, \mathbf{H} >$$

Prendre l'espérance ?

$$Q^{\mathcal{S}}(\mathcal{M}|\mathbf{D}) = \mathbb{E}_{\mathbf{H}\sim\mathbb{P}(\mathbf{H}|\mathbf{O},\mu)}\left[\mathcal{S}(\mathcal{M}|\mathbf{O},\mathbf{H})\right]$$

Mais la loi $\mathbb{P}(\mathbf{H}|\mathbf{O},\mu)$ est inconnue ! ! !

Donner un score à partir d'une base incomplète

Utiliser un modèle ?

Principe EM : E...

Supposons que $\mathcal{M}^0$ a généré **D**

$$\begin{aligned} Q^{\mathcal{S}}(\mathcal{M}|\mathbf{D}) &\approx Q^{\mathcal{S}}(\mathcal{M} : \mathcal{M}^0|\mathbf{D}) \\ &= \mathbb{E}_{\mathbf{H}\sim\mathbb{P}(\mathbf{H}|\mathbf{O},\mathcal{M}^0)}\left[\mathcal{S}(\mathcal{M}|\mathbf{O},\mathbf{H})\right] \\ &= \sum_{\mathbf{H}} \mathcal{S}(\mathcal{M}|\mathbf{O},\mathbf{H})\mathbb{P}(\mathbf{H}|\mathbf{O},\mathcal{M}^0) \end{aligned}$$

Ou la loi *a posteriori* $\mathbb{P}(\mathbf{H}|\mathbf{O},\mathcal{M}^0)$ est connue.

Structural-EM

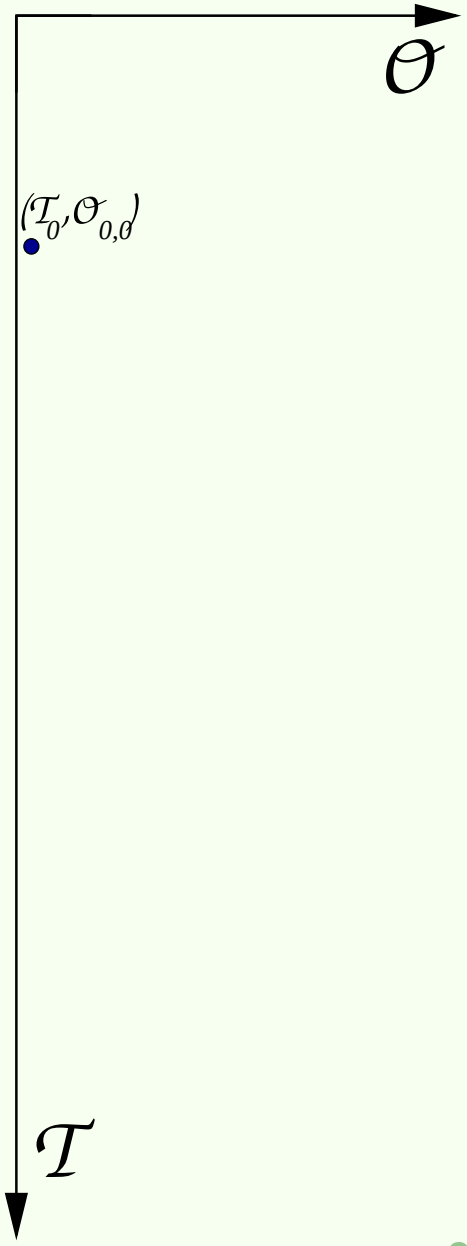
Principe EM : ...et M

- Choisir un modèle $\mathcal{M}^0$ ($\Rightarrow \mathbb{P}(\mathbf{H}|\mathbf{O}, \mathcal{M}^0)$)
- Trouver un modèle $\mathcal{M}^{i+1}$ qui *maximise*** le score $\mathcal{Q}^{\mathcal{S}}(\mathcal{M} : \mathcal{M}^i|\mathbf{D})$
- Utiliser le nouveau modèle comme référence pour l'itération suivante jusqu'à convergence.

** $\mathcal{M}^{i+1} = \{\mathcal{G}^{i+1}, \Theta^{i+1}\}$
Comment maximiser ?

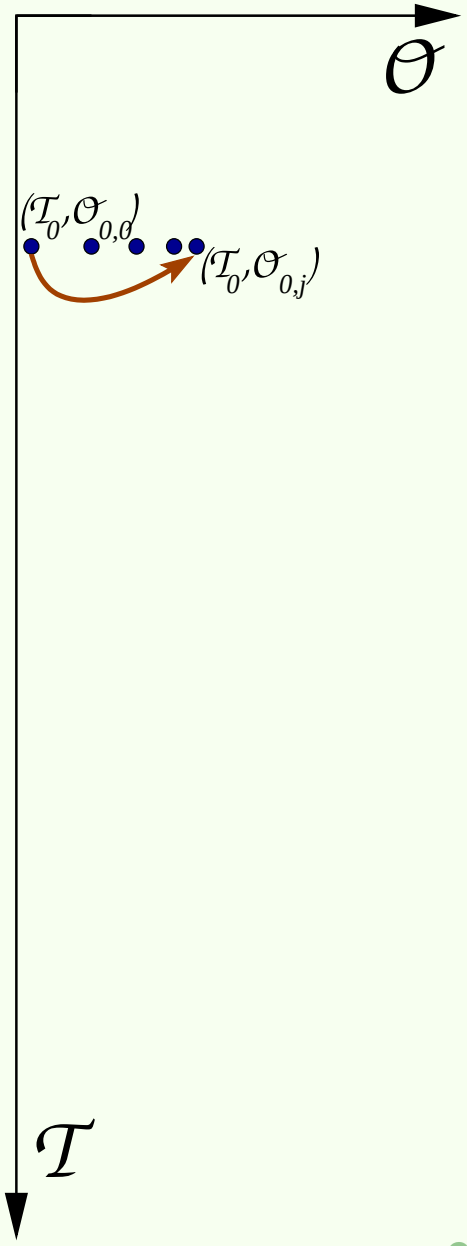
algorithme MWST-EM détaillé

1: **Init :** $fini = faux$, $i = 0$
Choix d'un réseau bayésien initial $(\mathcal{T}^0, \Theta^{0,0})$
2: **Répéter**



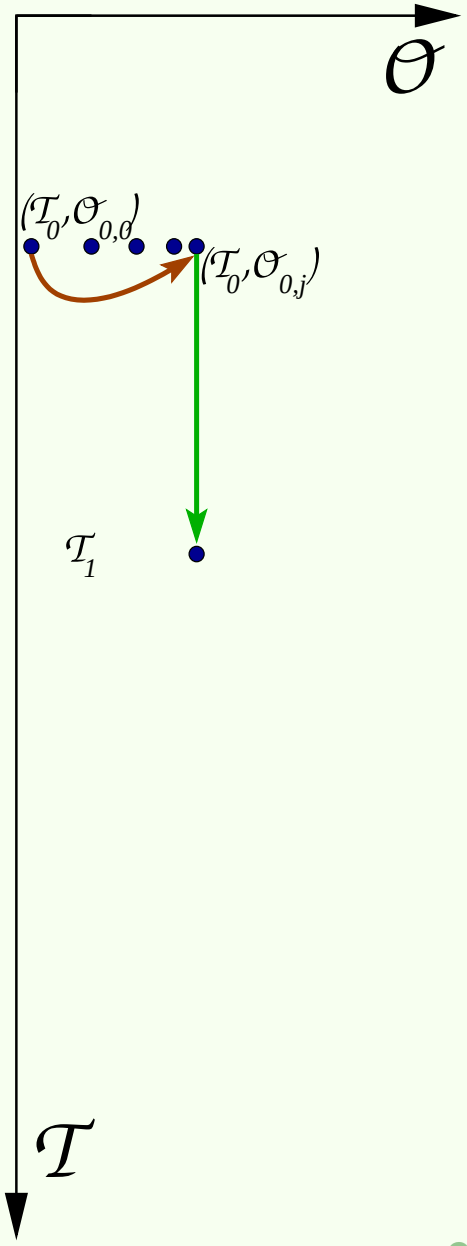
algorithme MWST-EM détaillé

1: **Init :** $fini = faux$, $i = 0$
Choix d'un réseau bayésien initial $(\mathcal{T}^0, \Theta^{0,0})$
2: **Répéter**
3: $j = 0$
4: **Répéter**
5: $\Theta^{i,j+1} = \underset{\Theta}{\operatorname{argmax}} \; Q(\mathcal{T}^i, \Theta : \mathcal{T}^i, \Theta^{i,j})$
6: $j = j + 1$
7: **Jusqu'à** convergence de $\Theta^{i,j}$



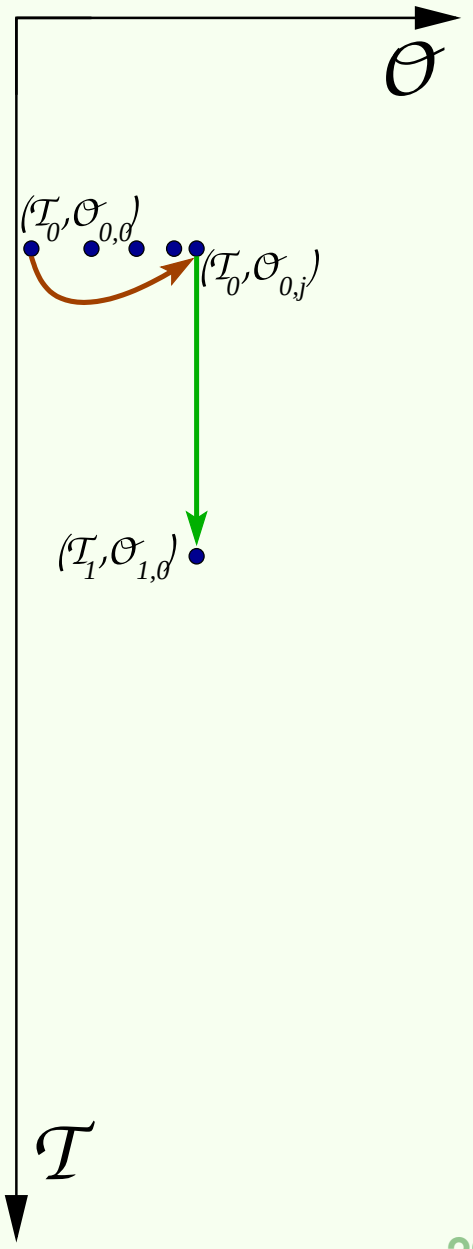
algorithme MWST-EM détaillé

1: **Init :** $fini = faux$, $i = 0$
Choix d'un réseau bayésien initial $(\mathcal{T}^0, \Theta^{0,0})$
2: **Répéter**
3: $j = 0$
4: **Répéter**
5: $\Theta^{i,j+1} = \underset{\Theta}{\operatorname{argmax}} \; Q(\mathcal{T}^i, \Theta : \mathcal{T}^i, \Theta^{i,j})$
6: $j = j + 1$
7: **Jusqu'à** convergence de $\Theta^{i,j}$
8: **Si** $i = 0$ ou $Q(\mathcal{T}^i, \Theta^{i,j} : \mathcal{T}^{i-1}, \Theta^{i-1,j}) - Q(\mathcal{T}^{i-1}, \Theta^{i-1,j} : \mathcal{T}^{i-1}, \Theta^{i-1,j}) > \epsilon$ **Alors**
9: $\mathcal{T}^{i+1} = \underset{\mathcal{T}}{\operatorname{argmax}} \; Q(\mathcal{T}, \bullet : \mathcal{T}^i, \Theta^{i,j})$
procédure de type **Kruskal**



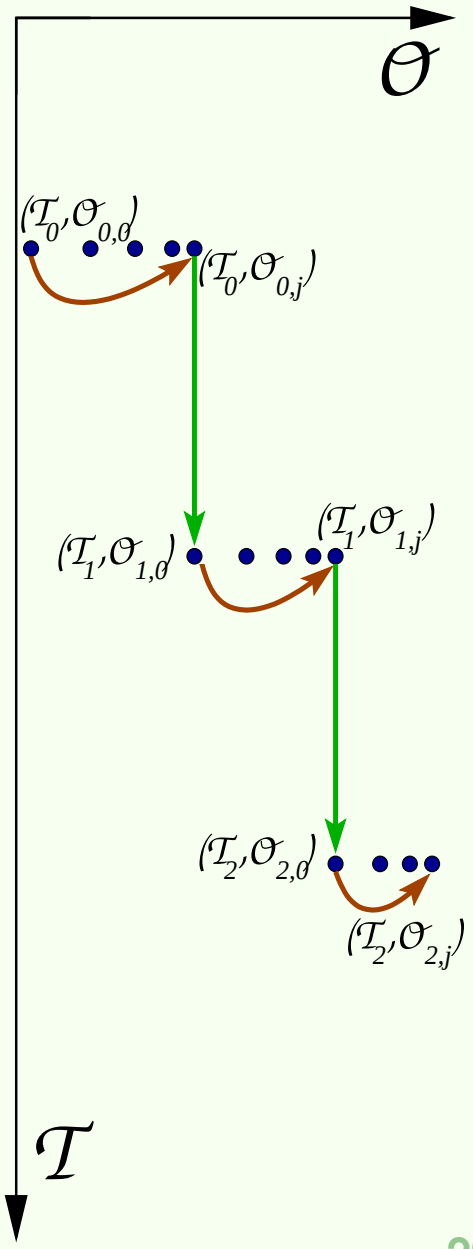
algorithme MWST-EM détaillé

1: **Init :** $fini = faux$, $i = 0$
Choix d'un réseau bayésien initial $(\mathcal{T}^0, \Theta^{0,0})$
2: **Répéter**
3: $j = 0$
4: **Répéter**
5: $\Theta^{i,j+1} = \underset{\Theta}{\operatorname{argmax}} \; Q(\mathcal{T}^i, \Theta : \mathcal{T}^i, \Theta^{i,j})$
6: $j = j + 1$
7: **Jusqu'à** convergence de $\Theta^{i,j}$
8: **Si** $i = 0$ ou $Q(\mathcal{T}^i, \Theta^{i,j} : \mathcal{T}^{i-1}, \Theta^{i-1,j}) - Q(\mathcal{T}^{i-1}, \Theta^{i-1,j} : \mathcal{T}^{i-1}, \Theta^{i-1,j}) > \epsilon$ **Alors**
9: $\mathcal{T}^{i+1} = \underset{\mathcal{T}}{\operatorname{argmax}} \; Q(\mathcal{T}, \bullet : \mathcal{T}^i, \Theta^{i,j})$
10: $\Theta^{i+1,0} = \underset{\Theta}{\operatorname{argmax}} \; Q(\mathcal{T}^{i+1}, \Theta : \mathcal{T}^i, \Theta^{i,j})$



algorithme MWST-EM détaillé

1: **Init :** $fini = faux$, $i = 0$
Choix d'un réseau bayésien initial $(\mathcal{T}^0, \Theta^{0,0})$
2: **Répéter**
3: $j = 0$
4: **Répéter**
5: $\Theta^{i,j+1} = \underset{\Theta}{\operatorname{argmax}} \; Q(\mathcal{T}^i, \Theta : \mathcal{T}^i, \Theta^{i,j})$
6: $j = j + 1$
7: **Jusqu'à** convergence de $\Theta^{i,j}$
8: **Si** $i = 0$ ou $Q(\mathcal{T}^i, \Theta^{i,j} : \mathcal{T}^{i-1}, \Theta^{i-1,j}) - Q(\mathcal{T}^{i-1}, \Theta^{i-1,j} : \mathcal{T}^{i-1}, \Theta^{i-1,j}) > \epsilon$ **Alors**
9: $\mathcal{T}^{i+1} = \underset{\mathcal{T}}{\operatorname{argmax}} \; Q(\mathcal{T}, \bullet : \mathcal{T}^i, \Theta^{i,j})$
10: $\Theta^{i+1,0} = \underset{\Theta}{\operatorname{argmax}} \; Q(\mathcal{T}^{i+1}, \Theta : \mathcal{T}^i, \Theta^{i,j})$
11: $i = i + 1$
12: **Sinon**
13: $fini = vrai$
14: **Fin Si**
15: **Jusqu'à** $fini = vrai$



Nos propositions MWST-EM et TAN-EM

SEM : le nouveau modèle est choisi parmis les *voisins* du graphe courant.

$\rightarrow$ nombreuses itérations

MWST-EM : nous trouvons le 'meilleur' modèle dans l'espace des arbres. (AKRR'05), (EGC'05)

$\rightarrow$ peu d'itérations

Utilisation d'une méthode de type '`Kruskal` itératif'

$$\left[M_{ij}^{Q}\right]_{i,j} = \left[Q^{bic}(X_i, P_i = \{X_j\} : \mathcal{T}^*) - Q^{bic}(X_i, P_i = \emptyset : \mathcal{T}^*)\right]$$

Nos propositions SEM+T et TAN-EM

SEM+T : le résultat de MWST-EM est utilisé pour initialiser SEM.

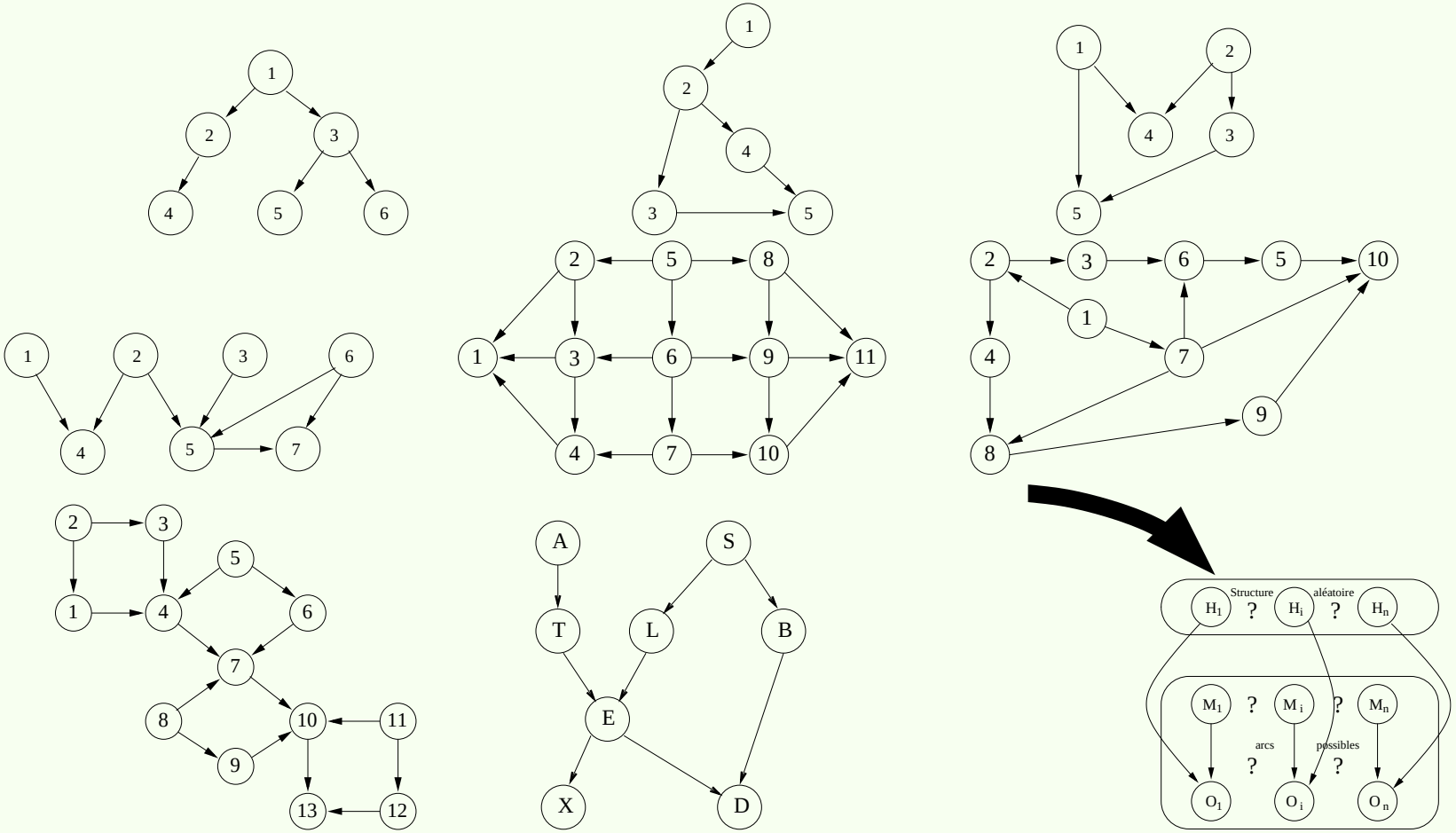
$\rightarrow$ nombreuses itérations ?

TAN-EM : nous trouvons le 'meilleur' Réseau Naïf Augmenté par un Arbre. (PGM'06)

$\rightarrow$ peu d'itérations

$$\left[M_{ij}^{Q}\right]_{i,j\neq\mathbf{C}} = \left[Q^{bic}(X_i, P_i = \{X_j, \mathbf{C}\} : \mathcal{T}^*) - Q^{bic}(X_i, P_i = \{\mathbf{C}\} : \mathcal{T}^*)\right]$$

Bases générés pour tester les méthodes



Taux de données manquantes : $\{0,2; 0,3; 0,4; 0,5\}$,
Tailles des bases d'exemples : $\{100; 200; 400; 600; 1000; 2000\}$.
$\longrightarrow$ 2160 problèmes MCAR synthétiques,
$\longrightarrow$ 2160 problèmes MAR sythétiques et 5 problèmes réels.

Protocol de tests

Nous avons testé :

- pertinence de EM *vs* CCA et ACA,
- performances BIC,
- performances KL,
- influence de la taille de la base d'exemples,
- influence du taux de données manquantes,
- stabilité BIC,
- stabilité KL,
- l'utilisation des RB en classification,
- temps de calcul.

sur des données MCAR et MAR synthétiques (2160 $\times$ 2 pbs)...

Protocol de tests

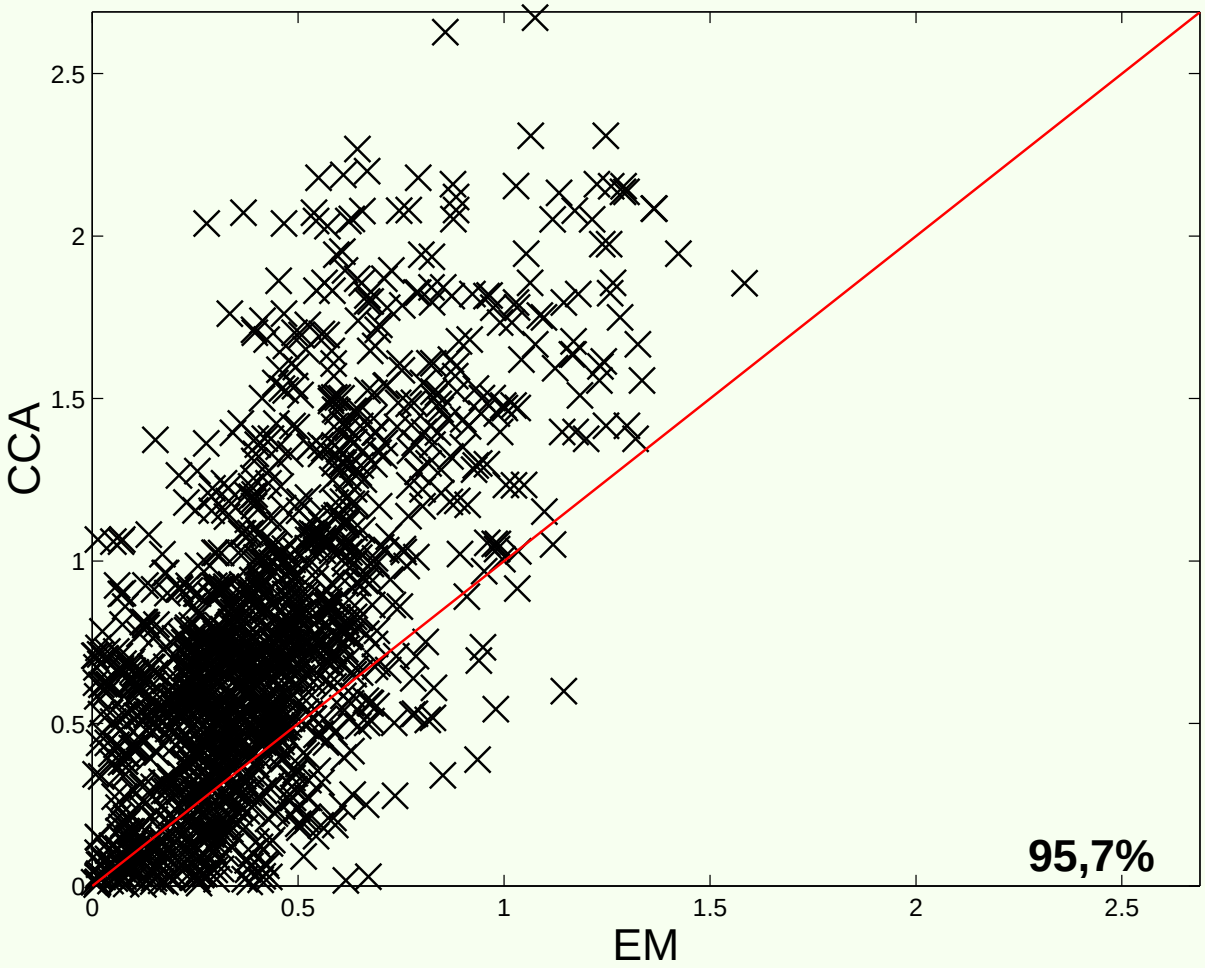
Nous avons testé :

- **pertinence de EM *vs* CCA et ACA**,
- performances BIC,
- **performances KL**,
- influence de la taille de la base d'exemples,
- **influence du taux de données manquantes**,
- stabilité BIC,
- stabilité KL,
- **l'utilisation des RB en classification**,
- **temps de calcul.**

sur des données MCAR et MAR synthétiques (2160 $\times$ 2 pbs)...

Pertinence de EM vs CCA et ACA

Comparaison de performances pour la divergence *KL* :

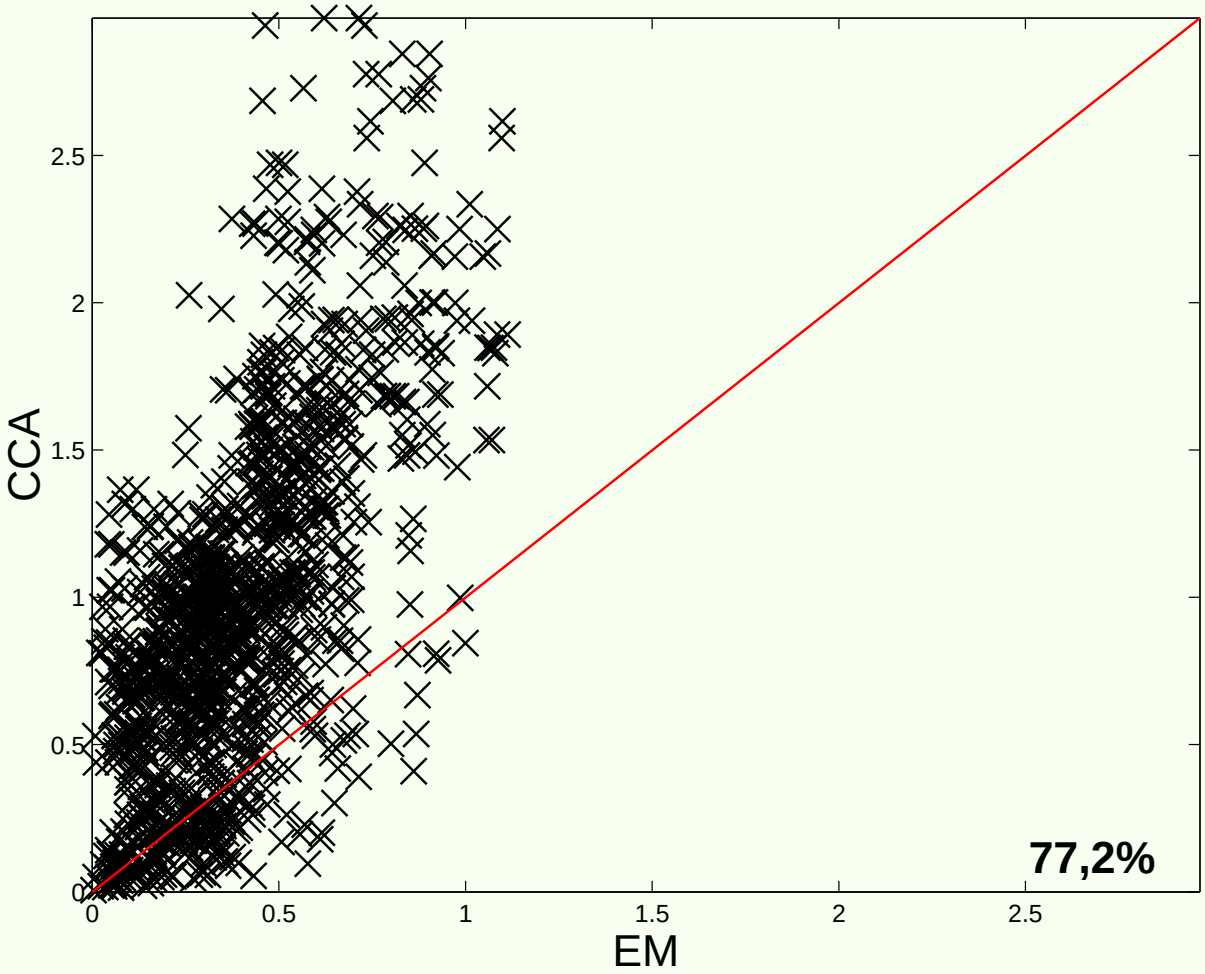


20% de données manquantes

MWST-CCA, GS-CCA, GS+T-CCA versus MWST-EM, SEM, SEM+T

Pertinence de EM vs CCA et ACA

Comparaison de performances pour la divergence *KL* :

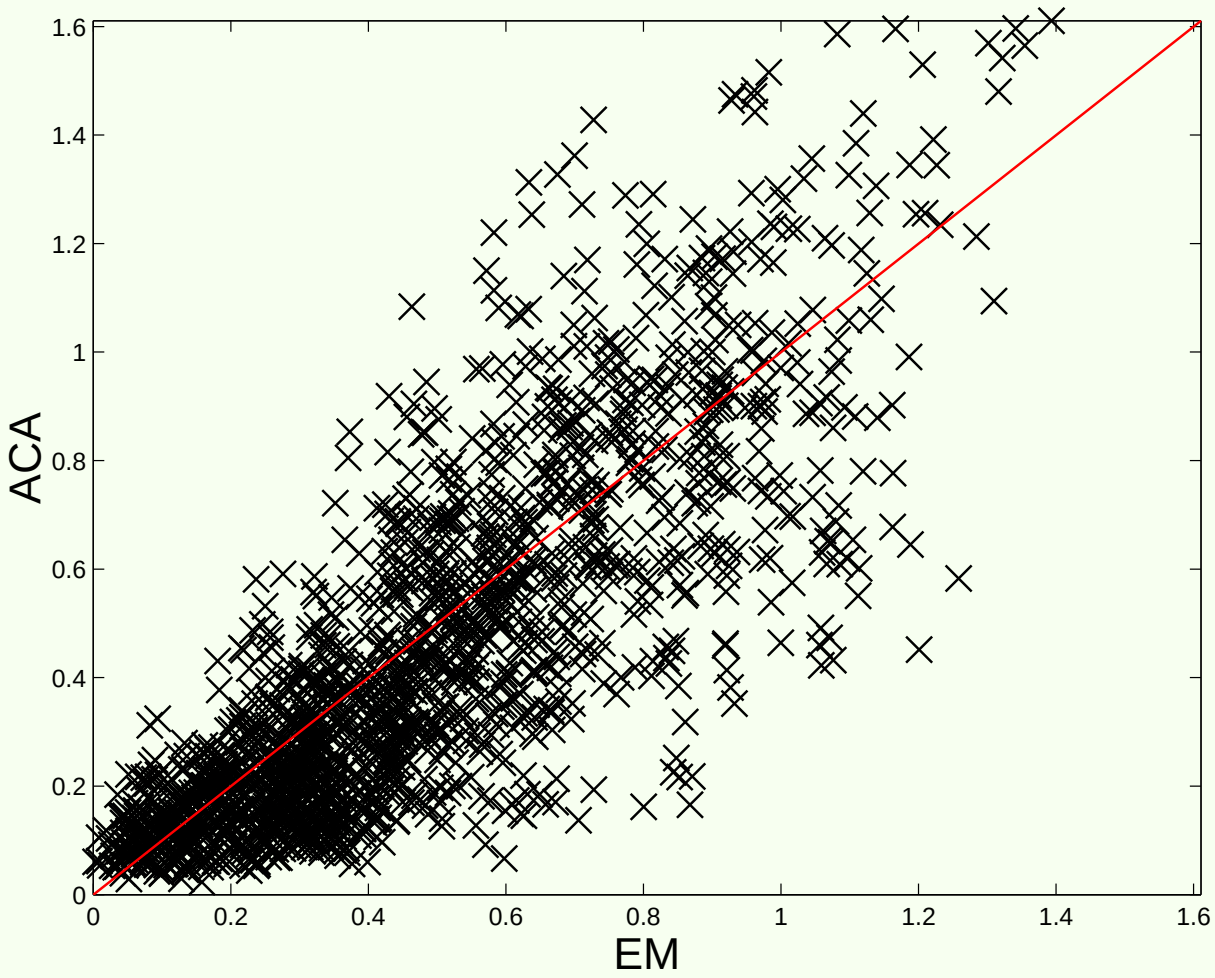


30% de données manquantes

MWST-CCA, GS-CCA, GS+T-CCA versus MWST-EM, SEM, SEM+T

Pertinence de EM vs CCA et ACA

Comparaison de performances pour la divergence *KL* :

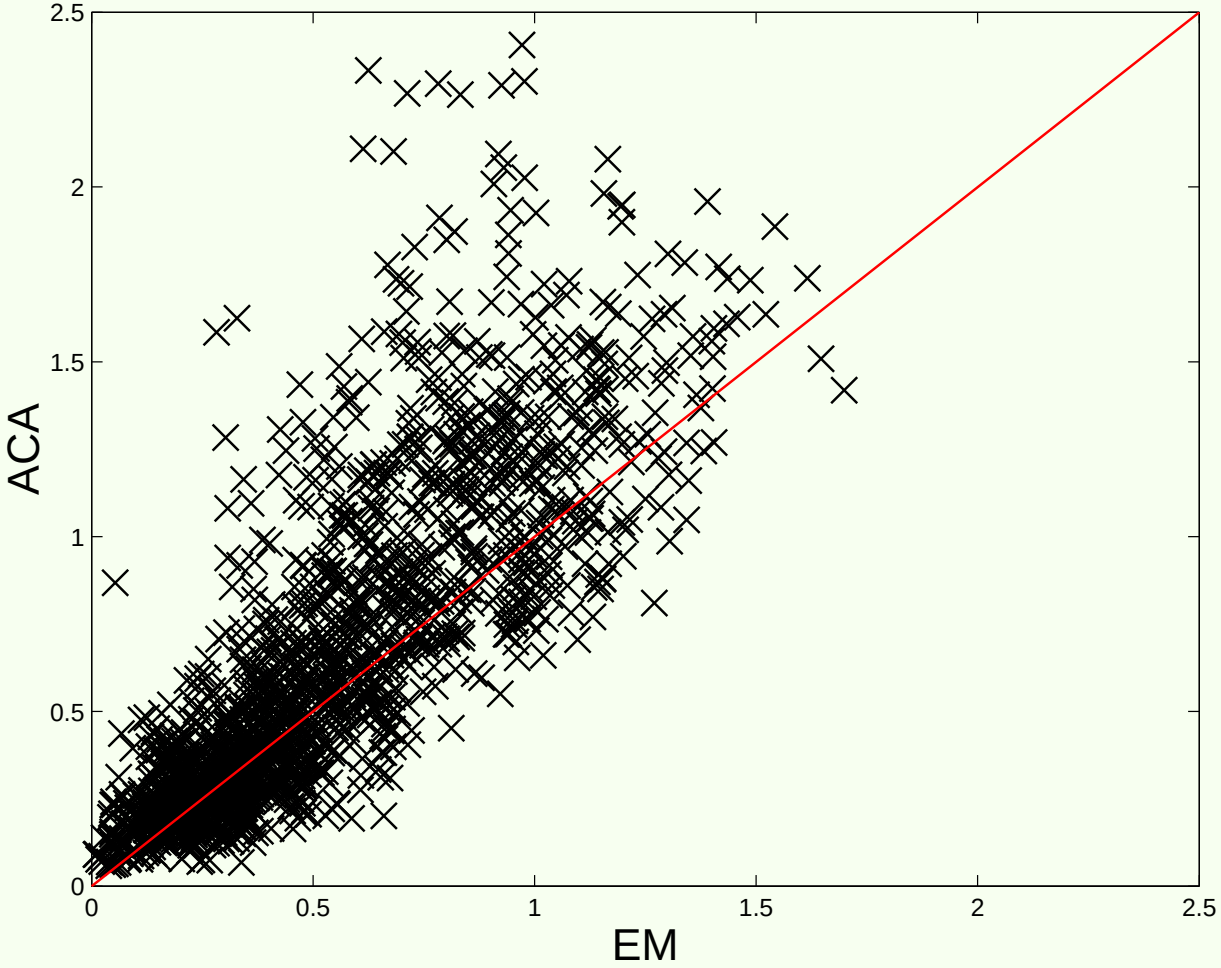


30% de données manquantes

MWST-ACA, GS-ACA, GS+T-ACA versus MWST-EM, SEM, SEM+T

Pertinence de EM vs CCA et ACA

Comparaison de performances pour la divergence *KL* :

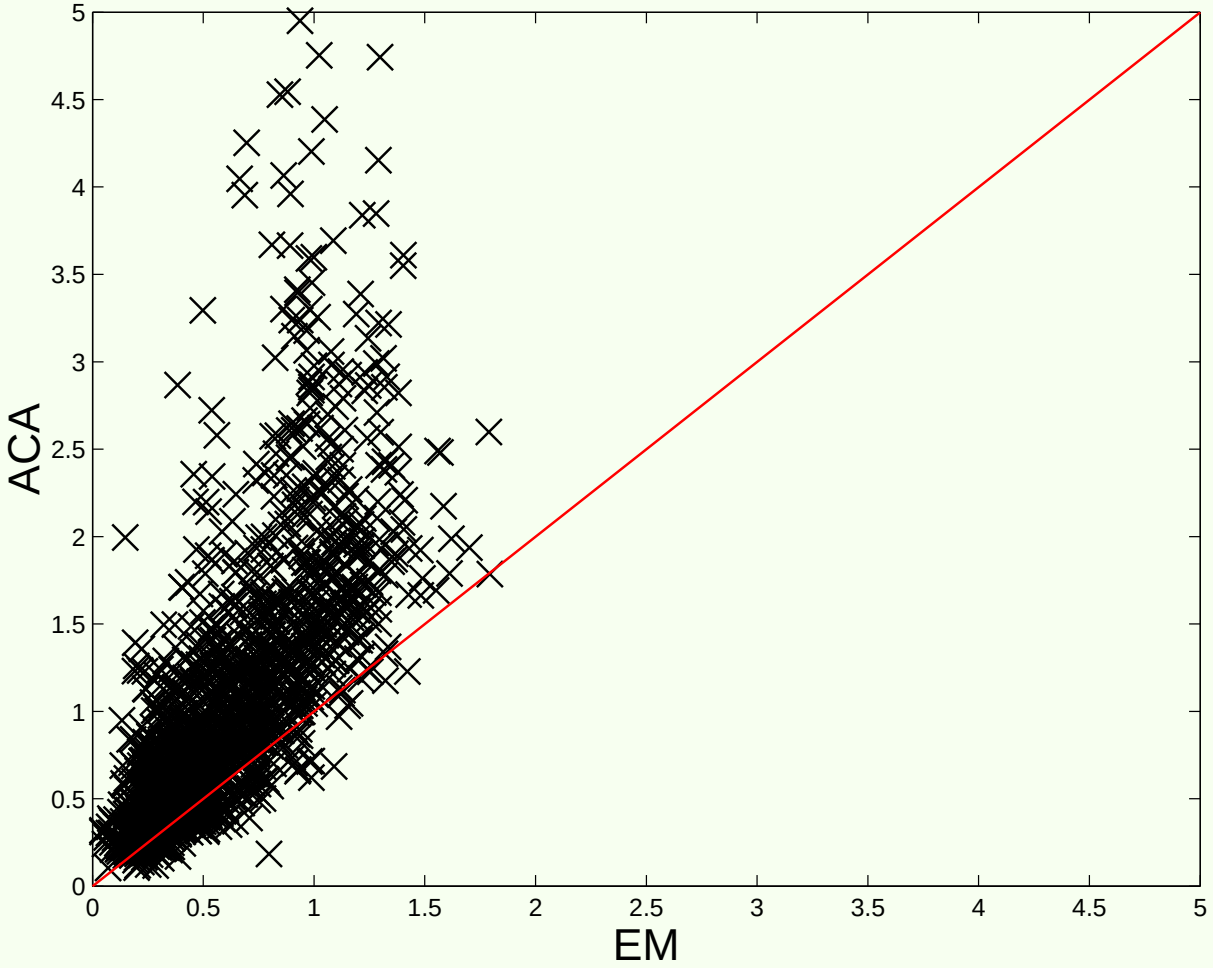


40% de données manquantes

MWST-ACA, GS-ACA, GS+T-ACA versus MWST-EM, SEM, SEM+T

Pertinence de EM vs CCA et ACA

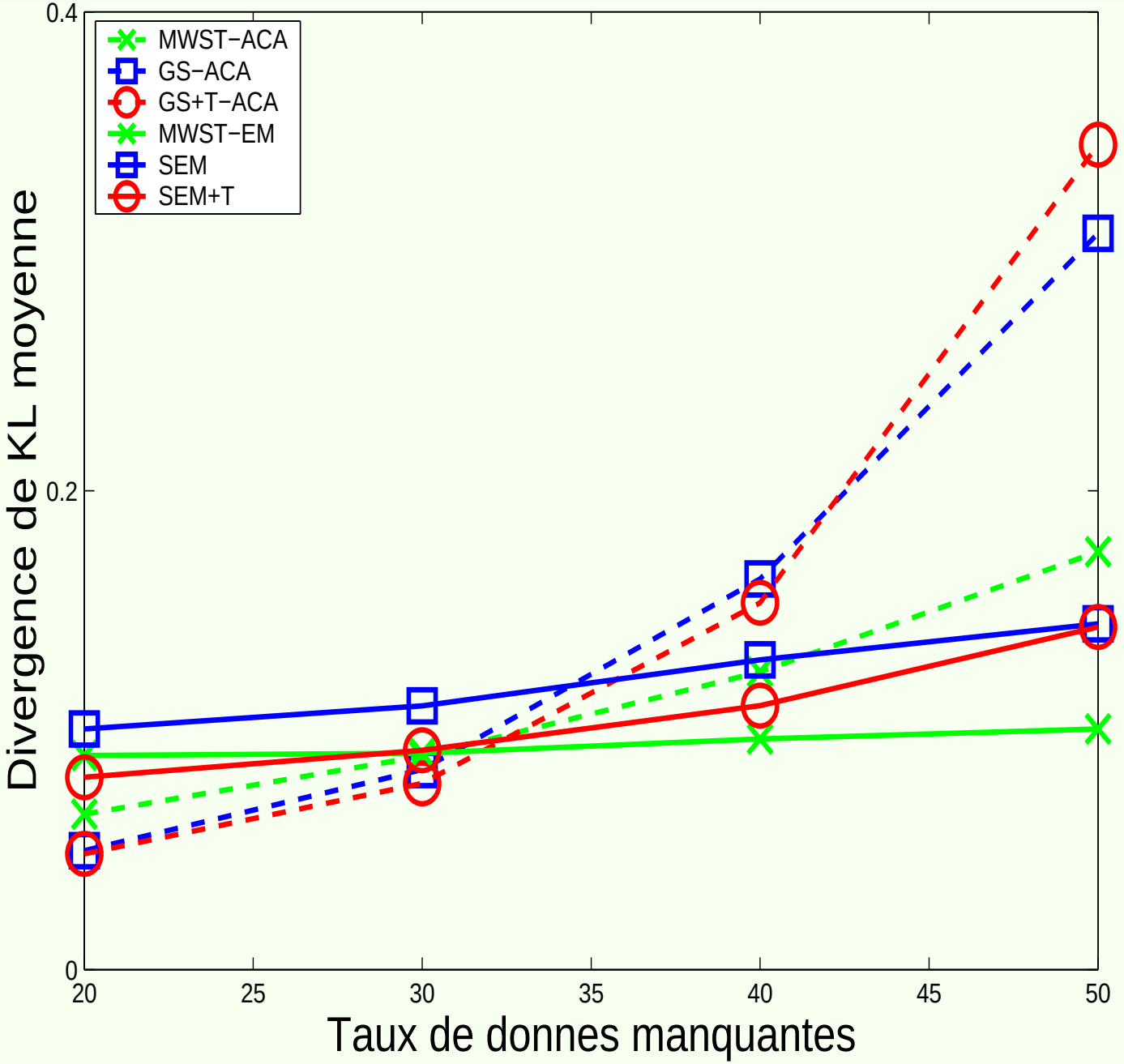
Comparaison de performances pour la divergence *KL* :



50% de données manquantes

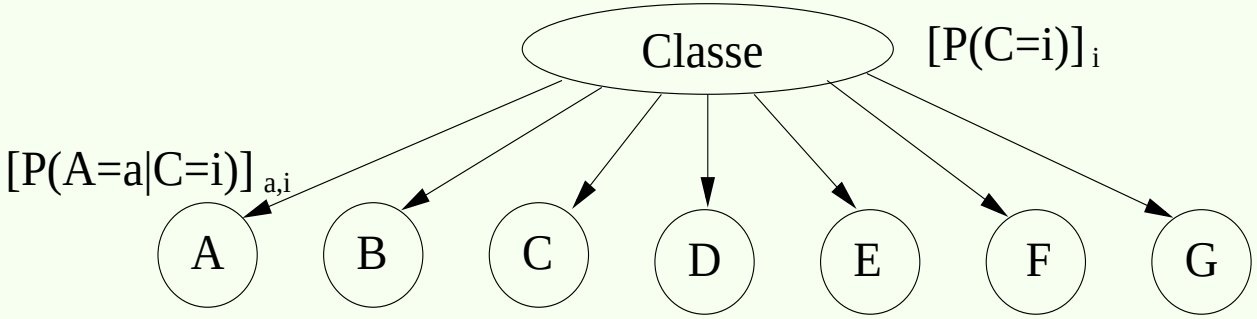
MWST-ACA, GS-ACA, GS+T-ACA versus MWST-EM, SEM, SEM+T

performances KL



Le réseau Naïf

Supposons que *la classe a une influence sur toutes les variables,* **mais** indépendamment

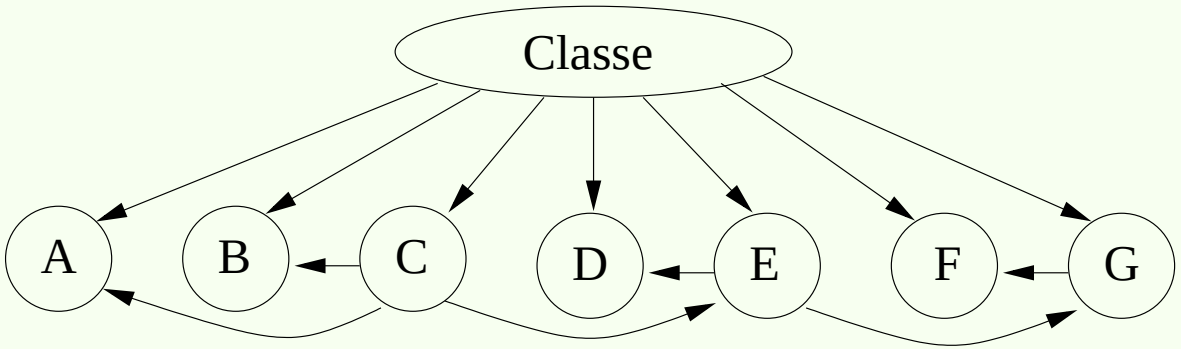


Ce qu'il est possible de faire :

- l'apprentissage des paramètres avec DI (par ex. avec EM),
- l'inférence avec des Données Incomplètes.

Le réseau Naïf

Supposons que *la classe a une influence sur toutes les variables*, **mais** *plus* indépendamment



Ce qu'il est *toujours* possible de faire :

- l'apprentissage des paramètres avec DI (par ex. avec EM),
- l'inférence avec des Données Incomplètes.

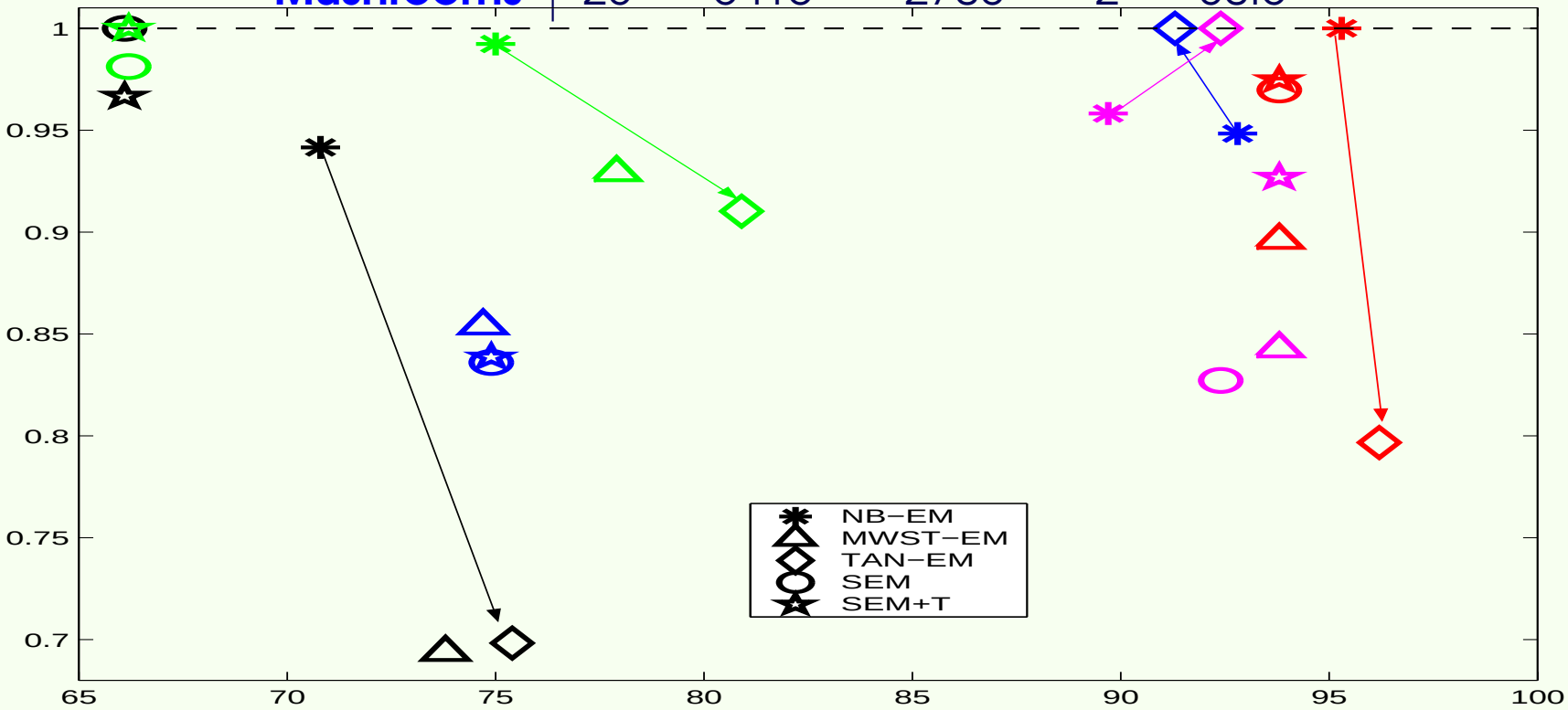
et

si l'on veut ajouter ces dépendances automatiquement ?

$\longrightarrow$ **TAN-EM**

Résultats en classification

	N	N app	N test	#C	%EI
Hepatitis	20	90	65	2	8.4
House	17	290	145	2	46.7
Horse	28	300	300	2	88.0
Thyroid	22	2800	972	2	29.9
Mushrooms	23	5416	2708	2	30.5



Temps de calculs

	NB-EM	MWST-EM	TAN-EM	SEM	SEM+T
tps de calcul	1,00	2,57	**2,56**	48,89	**33,00**
taux de classif 'moyens'	84,72	**82,80**	**87,24**	78,68	78,96

Conclusions

Nous avons **étudié empiriquement** les performances de l'**arbre couvrant de poids maximal** et **introduit** la méthode pour son **apprentissage** à partir de **bases d'exemples incomplètes**.

Bases d'exemples complètes ou incomplètes :

- Donne un **bon** résultat lorsque la **taille** de la base d'exemples est **faible**.
- Donne un **bon** résultat lorsque le **taux** de données manquantes est **élevé**.
- Est **très stable** par rapport à la **taille** de la base d'exemple.
- Est **très stable** par rapport au **taux** de données manquantes.

Conclusions

Nous avons **étudié empiriquement** la **pertinence d'initialiser** des méthodes gloutonnes avec un **arbre couvrant de poids maximal**.

Bases d'exemples complètes ou incomplètes :

- **Augmente** les **performances** et la **stabilité** de l'algorithme **K2**.
- **Augmente** la **stabilité** et **diminue** le **temps de calcul** de la méthode **GS** pour des performances équivalentes ou légèrement meilleures.

Conclusions

Nous avons **généralisé** la méthode **TAN** aux bases d'exemples incomplètes.

Bases d'exemples incomplètes :

- Très bonnes performances en classification.
- Temps d'apprentissage très court.

Conclusions

Nous avons **introduit** un formalisme de **modélisation** des **données manquantes**.

- Formalisme **génératif**.
- **Aisé** à comprendre pour les données **MCAR** et **MAR**.
- Complexe algorithmiquement lorsque le nombre d'attributs augmente ($\rightarrow$ inférence approchée).

...Et **diffusé** le code de toutes les fonctions utilisées.

Perspectives à court terme

Données Complètes

- **Tester** plus de méthodes (notamment les méthodes mixtes IC/score et celles d'identification de la frontière de Markov).
- **Mettre** en œuvre certaines méthodes pour les **bases mixtes**.
- **Adapter** les méthodes d'apprentissage à l'identification de la **causalité** entre les attributs,

Génération de bases

- **Mettre** en œuvre la méthode aux **données mixtes**.

Perspectives à court terme

Apprentissage avec données incomplètes

- (en cours) **Etendre MWST-EM** à la recherche de **forêts optimales** $\rightarrow$ Forest-EM.
- **Etendre** à la recherche de structure **optimale avec un nombre de parents fixé** (> 1).
- (en cours) **Etendre** SEM à la recherche dans l'espace des **équivalents de Markov** $\rightarrow$ GES-EM.
- (en cours) **Tester Forest-EM**
- (en cours) Tester **FAN-EM** en classification.
- **Généraliser TAN-EM** et FAN-EM au non-supervisé.

Perspectives à plus long terme

Apprentissage

- **Adapter** les méthodes à base de **tests statistiques** aux **bases incomplètes**.
- **Etendre** ces méthodes d'apprentissage aux **RBD**.
- **Etendre** les méthodes d'apprentissage aux **diagrammes d'influence**.
- (bientôt) **Etendre** les méthodes d'apprentissage aux POMDP (**processus de décision markoviens**).
- (bientôt) **Permettre** une **prise en compte d'information experte** plus simplement.
- **Permettre** un apprentissage **incrémental/adaptatif** pour les **bases non stationnaires**.
- **Passer** aux modèles partiellement orientés.

Revues :

Leray, P. et François, O. : *Réseaux Bayésiens pour la Classification - Méthodologie et Illustration dans le cadre du Diagnostic Médical*, **Revue d'Intelligence Artificielle**, ISBN : 2-7462-0912-8, vol. 18, n° 2/2004, pp169-193, 2004

François, O. et Leray, P. : *Etude Comparative d'Algorithmes d'Apprentissage de Structure dans les Réseaux Bayésiens*, **Journal électronique d'intelligence artificielle**, vol. 5, n° 39, pp1–19, 2004. Best paper de **RJCIA'03**.

Conférences internationales :

François, O.C.H et Leray, P. : *Learning the Tree Augmented Naive Bayes Classifier from incomplete datasets*, Proceedings of the Third European Workshop on Probabilistic Graphical Models (**PGM'06**), Prague, Czech Republic, ISBN : 80-86742-14-8, pp91-98, 2006.

Leray, P. et François, O. : *Bayesian Network Structural Learning and Incomplete Data*, International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning (**AKRR'05**), pp33-40, Finland, 2005.

Conférences nationales :

François, O. et Leray, P. : *Apprentissage de structure dans les réseaux bayésiens et données incomplètes*, journées Extraction et Gestion de Connaissances (Best paper de **EGC'05**), Revue des Nouvelles Technologies de l'Information (**RNTI-E-3**), Cépaduès Ed., ISBN : 2-85428-677-4, pp127–132, 2005,

François, O. et Leray, P. : *Evaluation d'algorithmes d'apprentissage de structure pour les réseaux bayésiens*, 14ieme Congrès francophone de Reconnaissance des formes et d'Intelligence artificielle (**RFIA'04**), pp1453-1460, 2004.

Autre publication :

Leray, P. et François, O. : *BNT Structure Learning Package : Documentation and Experiments*, Rapport technique du Laboratoire PSI, INSA de Rouen, n° 2004/PhLOF, 2004.
`http://bnt.insa-rouen.fr/`

Merci pour votre attention.

Questions ?
Remarques ?
Suggestions ?

Laboratoire d'Informatique, de Traitement de l'Information et des Systèmes

LITIS