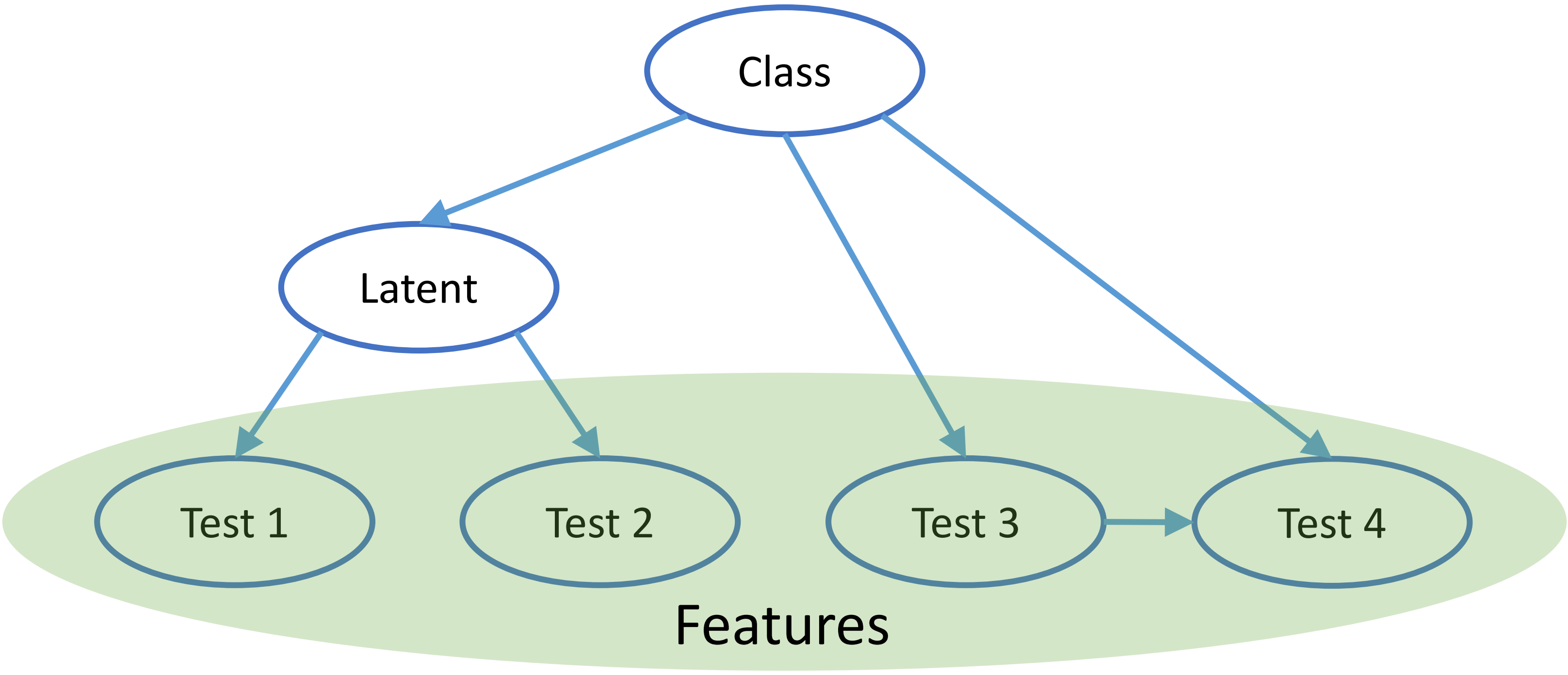
On Robust Trimming of Bayesian Network Classifiers

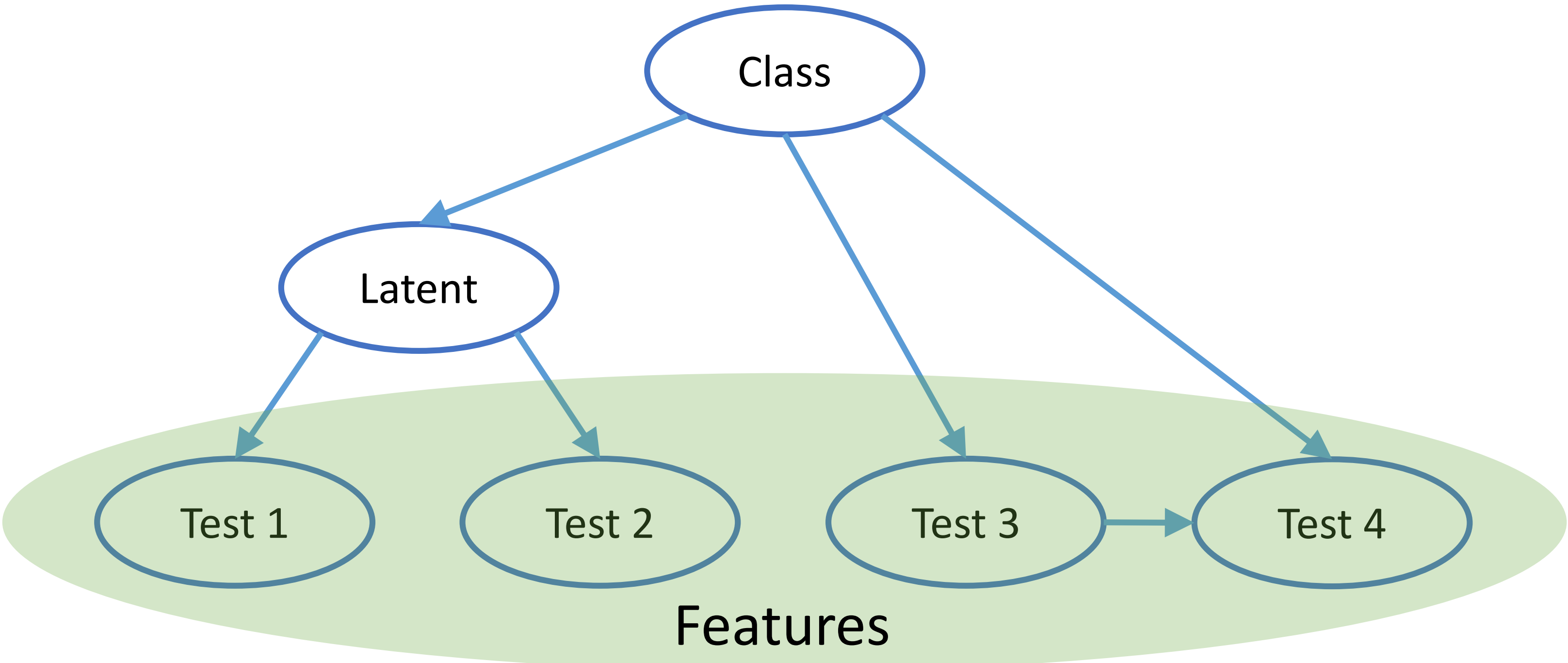
YooJung Choi and Guy Van den Broeck
UCLA

Bayesian Network Classifiers



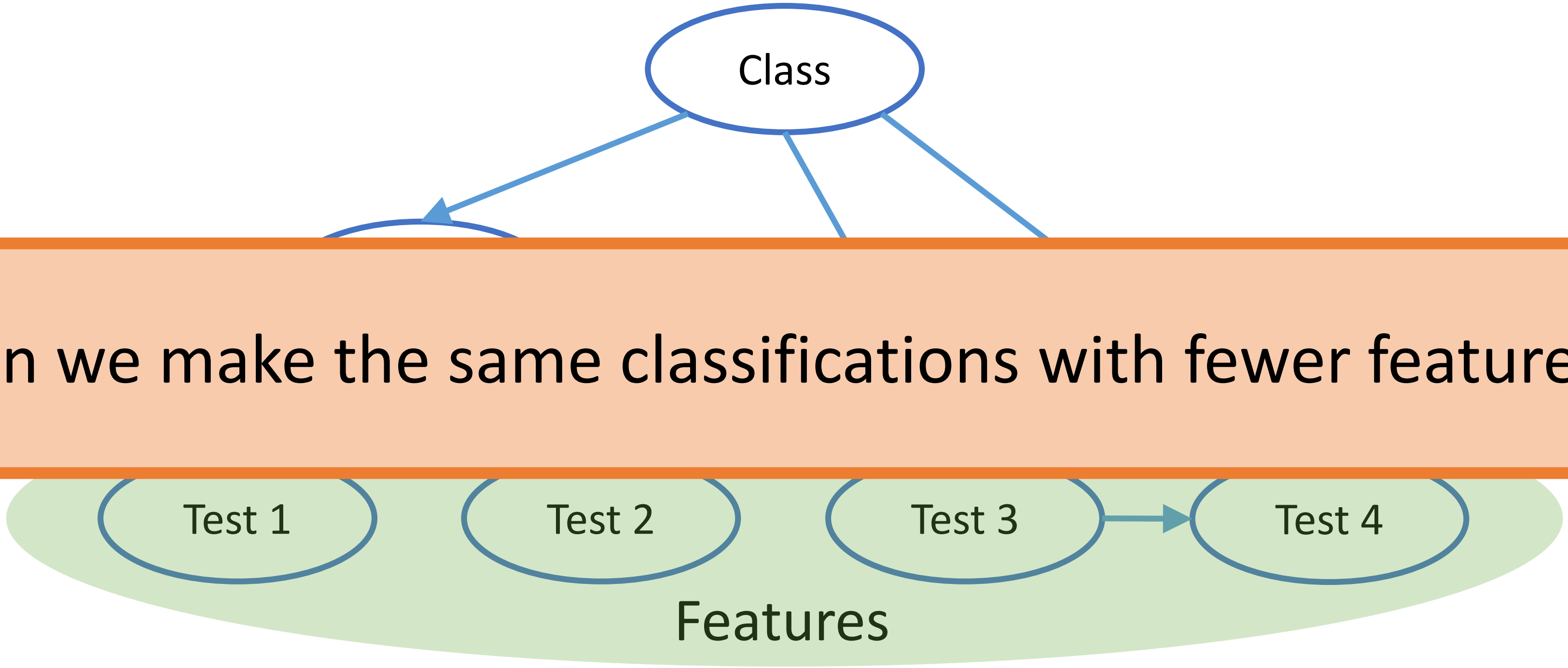
$$\Pr(C \mid \mathbf{features})$$

Bayesian Network Classifiers



$$C_T(\mathbf{features}) = \mathbb{I}\left(\Pr(C \mid \mathbf{features}) \geq T\right)$$

Bayesian Network Classifiers



Can we make the same classifications with fewer features?

$$C_T(\mathbf{features}) = \mathbb{I}\left(\Pr(C \mid \mathbf{features}) \geq T\right)$$

Why Classification Similarity?

To preserve classification behavior on individual examples

- Fairness
- Deployed classifiers

How to measure Similarity?

"Expected Classification Agreement"

$$\mathrm{ECA}(\alpha, \beta) = \sum_{\mathbf{f}} \mathbb{I}(C_T(\mathbf{f}) = C_{T'}(\mathbf{f}')) \cdot \Pr(\mathbf{f})$$

$\boxed{\text{What is the expected probability that a classifier } \alpha \text{ will agree with its trimming } \beta\text{?}}$

Robust Trimming

Original classifier

Trimmed classifier

$$\max_{\mathbf{F}' \subseteq \mathbf{F}} \max_{T'} \underbrace{\mathrm{ECA}(\alpha, (C, \mathbf{F}', T'))}_{\text{Similarity}}$$

$$\text{s.t. } \operatorname{cost}(\mathbf{F}') \leq B$$

Trimming Algorithm

Feature subset selection

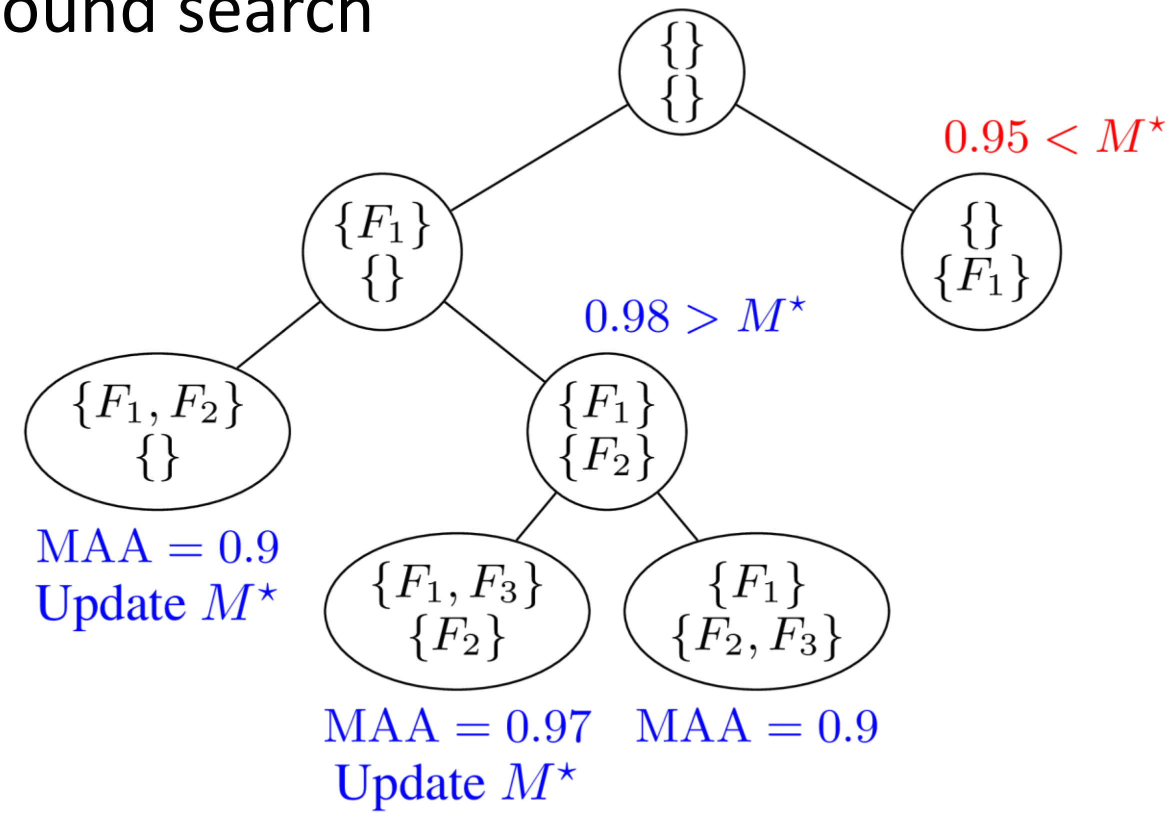
$$\max_{\mathbf{F}'\subseteq\mathbf{F}} \underbrace{\max_{T'} \mathrm{ECA}(\alpha,(C,\mathbf{F}',T'))}_{\text{“Maximum Achievable Agreement”}}$$

Search

Objective function

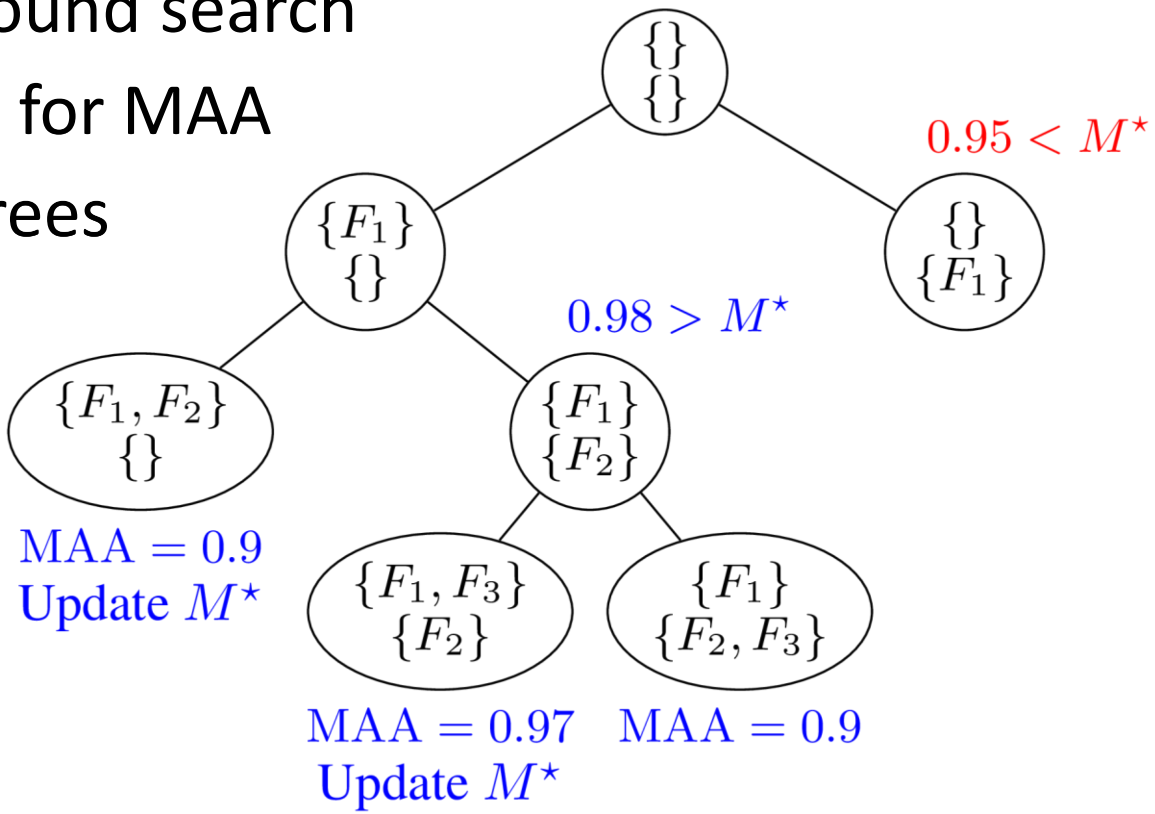
Trimming Algorithm

- Branch-and-Bound search



Trimming Algorithm

- Branch-and-Bound search
- Need a bound for MAA to prune subtrees



Upper-bound for MAA

"Maximum Potential Agreement"

$$\mathrm{MPA}_\alpha(\mathbf{F}') = \sum_{\mathbf{f}'} \max_c \sum_{\mathbf{f} \models \mathbf{f}'} \mathbb{I}(C_T(\mathbf{f}) = c) \Pr(\mathbf{f})$$

Maximum agreement between α and a hypothetical function that maps **f'** to c

Maximum Potential Agreement

1. Upper-bounds the MAA
2. Monotonically increasing

Great for pruning!

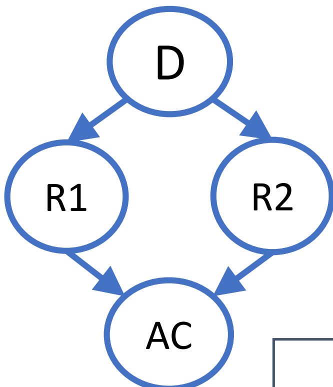
Maximum Potential Agreement

1. Upper-bounds the MAA
2. Monotonically increasing
3. Generally easier to compute than MAA
4. Equal to MAA given some independence condition (e.g. Naïve Bayes)

Great for pruning!

Computing the MPA and MAA

Prior works based on knowledge compilation



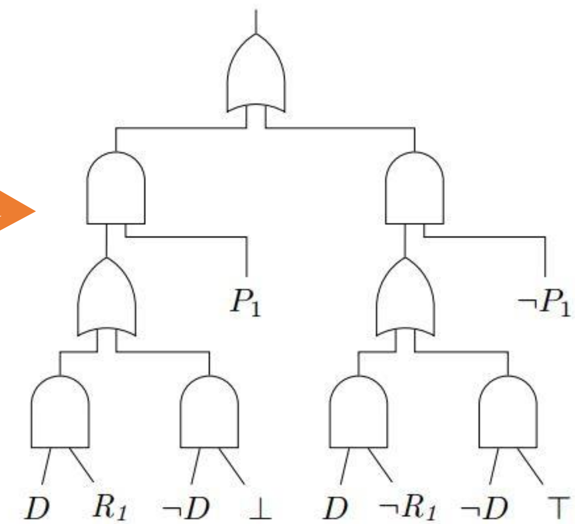
D	$\Pr(R_1 = + \mid D)$
$+$	0.7
$-$	0.2

$\Pr(D = +)$
0.2

…

$$P_1 \Leftrightarrow D \land R_1$$
$$P_2 \Leftrightarrow D \land \neg R_1$$
$$P_3 \Leftrightarrow \neg D \land R_1$$
$$P_4 \Leftrightarrow \neg D \land \neg R_1$$

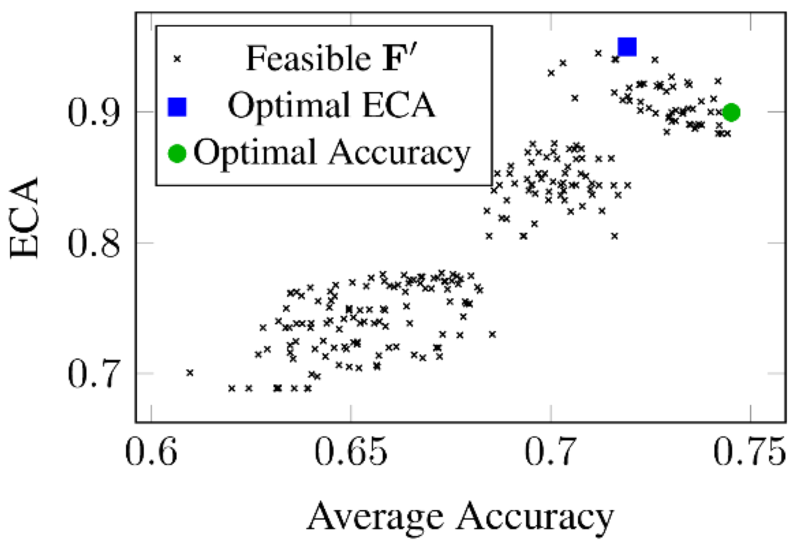
$$w(P_1) = 0.7$$
$$w(P_2) = 0.3$$
$$w(P_3) = 0.2$$
$$w(P_4) = 0.8$$
$$w(l) = 1.0 \text{ for all other literal } l$$

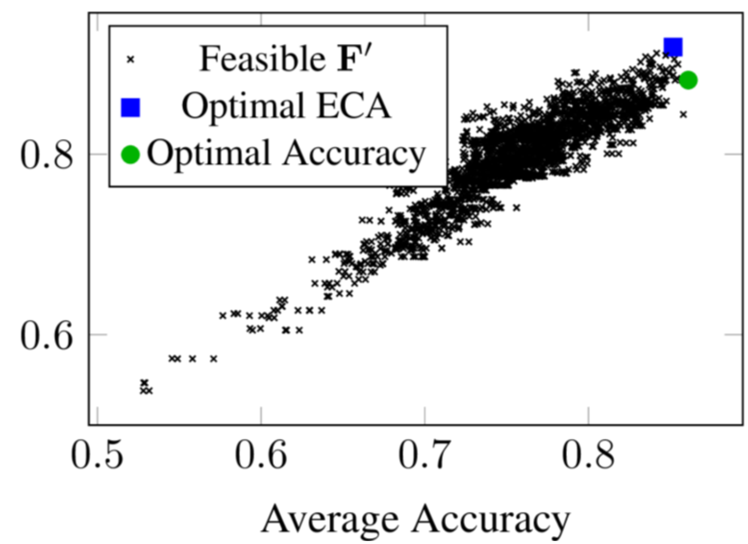


(a) Sentential decision diagram (SDD)

[Oztok,Choi,Darwiche 2016; C,Darwiche,VdB 2017]

Evaluation





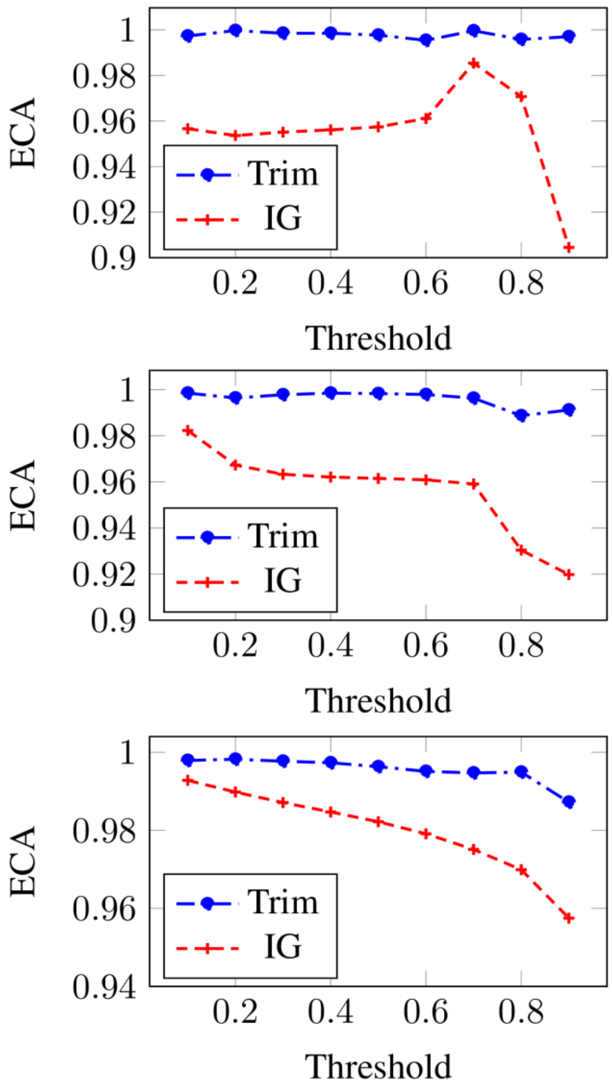
		Agreement	Accuracy
`pima`	Opt. ECA	0.9863	0.7123
	Opt. Acc.	0.9452	0.7260
`heart`	Opt. ECA	0.9245	0.8491
	Opt. Acc.	0.9057	0.7925

Evaluation

	$\lvert\mathbf{F}\rvert$	FS-SDD Time	FS-SDD # Eval	ECA-TRIM Time	ECA-TRIM # Eval	$\binom{n}{m}$
bupa	6	0.044	21	0.026	14	15
pima	8	0.056	36	0.039	45	28
ident	9	0.128	129	0.097	89	84
anatomy	12	2.252	793	1.085	283	495
heart	13	7.346	1092	2.234	209	715
voting	16	819.163	6884	407.571	3345	4368
hepatitis	19	Timeout	43795	4390.71	2208	27132

Branch-and-bound improves efficiency (even with extra upper-bound computations)

Evaluation



High information gain does not lead to high classification agreement

Information-theoretic measures unaware of changes in classification threshold

Thank you!

Questions?