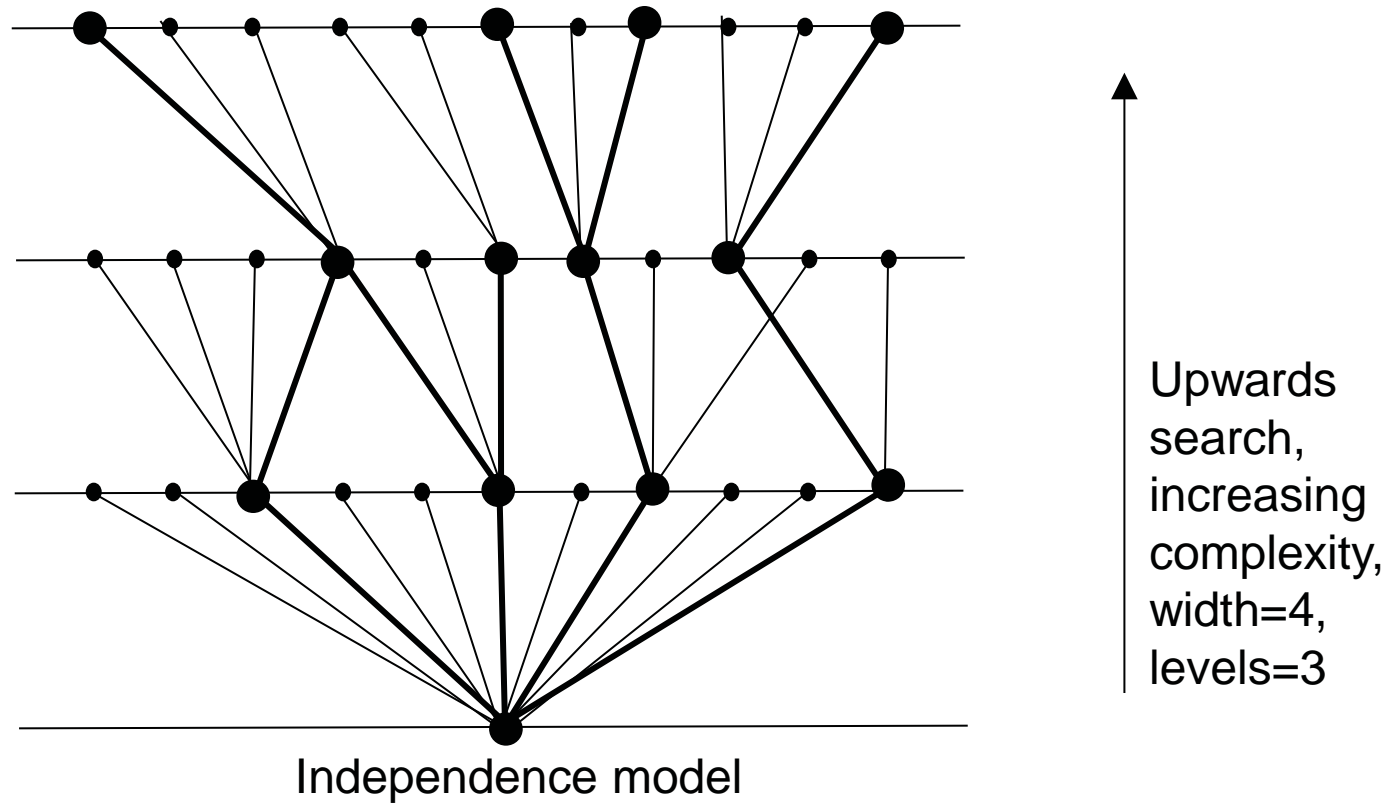


OCCAM search output (1/4)

- Search many models: select best w ('width') models at each of l 'levels'
- Output all selected $w \cdot l$ models
- OCCAM summarizes the best of these

OCCAM search output (2/4)



OCCAM search output (3/4)

- OCCAM summarizes:
- Of width*levels models found in search, it indicates the **best three** by different criteria:
 - (i) highest Δ **BIC**
 - (ii) highest Δ **AIC**
 - (iii) highest information with OK **p-values**
(both **cumulative** & **incremental**)

Sample input file: dementia05.txt (for demo#2,#3)

```
#Fifth test set of OHSU data 20 June 2012
#Corrected transform of variable I rs12248379
#Corrected transform of APOE
#Updated imputed data
#Deleted var I and dropped records with no var C data
#Added stats on missing. Blank = none missing
```

```
:nominal
ID ,0,0,ID
APOE ,2,1,AP
Gender ,2,1,SX
Education ,3,1,ED
AgeLastExam ,3,1,AG
rs1801133 ,3,1,A
rs3818361 ,4,1,B # missing 3
rs7561528 ,3,1,C
rs744373 ,3,1,D
rs6943822 ,3,1,E
rs4298437 ,3,1,F
rs7012010 ,3,1,G
rs11136000 ,3,1,H
rs10786998 ,4,1,J # missing 9
rs11193130 ,4,1,K # missing 11
rs610932 ,3,1,L
rs3851179 ,3,1,M
rs3764650 ,4,1,N # missing 2
rs3865444 ,4,1,P # missing 9
CaseControl ,2,2,Z
```

```
:no-frequency
```

```
:data
#IndID APOE GENDER EDU ALE A B C D E F G H J
101 0 0 2 2 1 1 0 1 2 2 1 1 2
103 0 0 2 1 0 2 2 0 1 1 1 2 2
111 0 1 2 1 2 2 1 1 0 1 1 2 1
112 0 0 2 2 2 2 1 1 1 2 1 1 0
118 0 1 0 2 2 2 2 0 0 1 1 1 .
120 0 1 2 2 1 2 1 1 0 1 1 2 1
121 0 0 2 2 2 2 1 1 2 0 0 0 2
122 0 0 1 2 1 2 1 1 2 0 0 2 2
123 0 0 2 2 2 2 2 0 1 1 0 0 2
126 0 0 2 2 2 2 2 0 1 1 2 2 2
127 0 1 2 2 0 2 1 1 1 1 1 2 2
128 0 0 2 2 1 2 1 1 0 1 1 2 .
129 1 1 2 2 0 2 1 0 1 0 1 2 0
132 0 0 2 2 2 2 1 0 0 1 1 2 1
134 0 1 2 2 0 0 2 0 0 1 1 1 0
135 0 0 2 2 2 1 1 0 0 2 1 1 0
```

Search output for dementia05.txt

VB models with loops (search sort on info)(demo#2)

ID	MODEL	Level	H	dDF	dLR	Alpha	Inf	%dH(DV)	dAIC	dBIC	Inc.Alpha	Prog.	%C(Data)	%cover
22	IV:ApZ:EdCZ:EdKZ:LZ	7	9.5221	20	120.2026	0.0000	0.20476589	20.4766	80.2026	-0.7921	0.0505	19	72.4057	52.7778
21	IV:ApZ:EdCZ:EdKZ:BZ	7	9.5224	21	120.0402	0.0000	0.20448919	20.4489	78.0402	-7.0043	0.1093	18	72.6415	38.5417
20	IV:ApZ:EdKZ:EdLZ:CZ	7	9.5224	20	120.0236	0.0000	0.20446093	20.4461	80.0236	-0.9711	0.0890	17	71.9340	52.7778
19	IV:ApZ:EdCZ:EdKZ	6	9.5323	18	114.2291	0.0000	0.19459005	19.4590	78.2291	5.3339	0.1469	16	71.6981	76.3889
18	IV:ApZ:EdKZ:BZ:CZ	6	9.5352	17	112.5147	0.0000	0.19166949	19.1669	78.5147	9.6692	0.1656	16	72.1698	38.5417
17	IV:ApZ:EdKZ:CZ:LZ	6	9.5361	16	111.9806	0.0000	0.19075974	19.0760	79.9806	15.1849	0.0053	14	72.4057	52.7778
16*	IV:ApZ:EdKZ:CZ	5	9.5438	14	107.4549	0.0000	0.18305008	18.3050	79.4549	22.7586	0.0037	13	71.2264	76.3889
15*	IV:ApZ:EdJZ:CZ	5	9.5536	14	101.7185	0.0000	0.17327808	17.3278	73.7185	17.0222	0.0434	11	70.7547	76.3889
14	IV:ApZ:EdKZ:LZ	5	9.5540	14	101.4868	0.0000	0.17288344	17.2883	73.4868	16.7905	0.0737	13	71.6981	76.3889
13*	IV:ApZ:EdKZ	4	9.5628	12	96.2708	0.0000	0.16399792	16.3998	72.2708	23.6740	0.0274	10	70.5189	91.6667
12*	IV:ApZ:EdZ:CZ:KZ	4	9.5665	8	94.1440	0.0000	0.16037485	16.0375	78.1440	45.7461	0.0007	8	70.0472	76.3889
11*	IV:ApZ:EdZ:CZ:JZ	4	9.5756	8	88.7544	0.0000	0.15119366	15.1194	72.7544	40.3565	0.0301	9	69.5755	76.3889
10*	IV:ApZ:EdZ:KZ	3	9.5870	6	82.0689	0.0000	0.13980489	13.9805	70.0689	45.7705	0.0003	5	69.8113	91.6667
9*	IV:ApZ:EdZ:CZ	3	9.5908	5	79.8373	0.0000	0.13600335	13.6003	69.8373	49.5886	0.0010	6	69.5755	100.0000
8*	IV:ApZ:CZ:KZ	3	9.5910	6	79.7227	0.0000	0.13580807	13.5808	67.7227	43.4243	0.0009	5	68.6321	87.5000
7*	IV:ApZ:EdZ	2	9.6133	3	66.5901	0.0000	0.11343675	11.3437	60.5901	48.4409	0.0000	2	69.5755	100.0000
6*	IV:ApZ:CZ	2	9.6144	3	65.9506	0.0000	0.11234732	11.2347	59.9506	47.8014	0.0000	3	67.9245	100.0000
5*	IV:ApZ:KZ	2	9.6150	4	65.6244	0.0000	0.11179159	11.1792	57.6244	41.4255	0.0024	4	66.9811	100.0000
4*	IV:ApZ	1	9.6396	1	51.1652	0.0000	0.08716028	8.7160	49.1652	45.1155	0.0000	1	66.9811	100.0000
3*	IV:CZ	1	9.6998	2	15.7735	0.0004	0.02687026	2.6870	11.7735	3.6740	0.0004	1	58.0189	100.0000
2*	IV:EdZ	1	9.7009	2	15.1356	0.0005	0.02578359	2.5784	11.1356	3.0361	0.0005	1	56.3679	100.0000
1*	IV:Z	0	9.7266	0	0.0000	1.0000	0.00000000	0.0000	0.0000	0.0000	0.0000	0	52.1226	100.0000
ID	MODEL	Level	H	dDF	dLR	Alpha	Inf	%dH(DV)	dAIC	dBIC	Inc.Alpha	Prog.	%C(Data)	%cover

Best Model(s) by dBIC:

9*	IV:ApZ:EdZ:CZ	3	9.5908	5	79.8373	0.0000	0.13600335	13.6003	69.8373	49.5886	0.0010	6	69.5755	100.0000
----	---------------	---	--------	---	---------	--------	------------	---------	---------	---------	--------	---	---------	----------

Best Model(s) by dAIC:

22	IV:ApZ:EdCZ:EdKZ:LZ	7	9.5221	20	120.2026	0.0000	0.20476589	20.4766	80.2026	-0.7921	0.0505	19	72.4057	52.7778
----	---------------------	---	--------	----	----------	--------	------------	---------	---------	---------	--------	----	---------	---------

Best Model(s) by Information, with all Inc. Alpha < 0.05:

16*	IV:ApZ:EdKZ:CZ	5	9.5438	14	107.4549	0.0000	0.18305008	18.3050	79.4549	22.7586	0.0037	13	71.2264	76.3889
-----	----------------	---	--------	----	----------	--------	------------	---------	---------	---------	--------	----	---------	---------

Search output (4/4)

Goodness measures for best models selected using BIC, AIC, p-value criteria:

- Reduction of DV uncertainty $\% \Delta H(DV | IV)$,
- Model complexity, df (degrees of freedom)
- Predictive accuracy: %correct

Search output for dementia05.txt

Models selected by BIC/AIC/p-value (sort on [info](#))

<u>Criterion</u>	<u>model</u>	<u>%ΔH</u>	<u>df</u>	<u>%c</u>
BIC	IV:ApZ:EdZ:CZ	14	5	70
p-value	IV:ApZ:EdKZ:CZ	18	14	71
AIC	IV:ApZ:EdCZ:EdKZ:LZ	20	20	72
SB BIC	IV:Ap ₀ Ed ₀ Z:Ap ₁ Z:C ₂ Z:Z	14	3	70

IV = ApEdCKL... (all independent variables)

Fit output for dementia05.txt

model IV:ApZ:EdZ:CZ (has loop) (demo#3)

Conditional DV (D) (%) for each IV composite state for the Model IV:ApZ:EdZ:CZ.
IV order: ApEdC (APOE; Education; rs7561528).

IV			Data			Model							
Ap	Ed	C	freq	obs. p(DV IV)		calc. q(DV IV)			rule	#correct	%correct	p(rule)	p(margin)
0	0	0	6.000	33.333	66.667	16.719	83.281	1	4.000	66.667	0.103	0.083	
0	0	1	4.000	0.000	100.000	21.075	78.925	1	4.000	100.000	0.247	0.214	
0	0	2	7.000	14.286	85.714	35.284	64.716	1	6.000	85.714	0.436	0.372	
0	1	0	8.000	50.000	50.000	52.532	47.468	0	4.000	50.000	0.886	0.981	
0	1	1	40.000	62.500	37.500	59.546	40.454	0	25.000	62.500	0.227	0.347	
0	1	2	44.000	72.727	27.273	75.034	24.966	0	32.000	72.727	0.001	0.002	
0	2	0	14.000	42.857	57.143	55.306	44.694	0	6.000	42.857	0.691	0.811	
0	2	1	86.000	60.465	39.535	62.205	37.795	0	52.000	60.465	0.023	0.061	
0	2	2	68.000	83.824	16.176	77.067	22.933	0	57.000	83.824	0.000	0.000	
1	0	0	3.000	0.000	100.000	3.984	96.016	1	3.000	100.000	0.111	0.095	
1	0	1	5.000	20.000	80.000	5.230	94.770	1	4.000	80.000	0.045	0.036	
1	0	2	3.000	33.333	66.667	10.127	89.873	1	2.000	66.667	0.167	0.146	
1	1	0	8.000	25.000	75.000	18.614	81.386	1	6.000	75.000	0.076	0.058	
1	1	1	19.000	26.316	73.684	23.325	76.675	1	14.000	73.684	0.020	0.012	
1	1	2	21.000	33.333	66.667	38.315	61.685	1	14.000	66.667	0.284	0.206	
1	2	0	12.000	25.000	75.000	20.366	79.634	1	9.000	75.000	0.040	0.028	
1	2	1	40.000	25.000	75.000	25.381	74.619	1	30.000	75.000	0.002	0.001	
1	2	2	36.000	36.111	63.889	40.986	59.014	1	23.000	63.889	0.280	0.181	
			424.000	52.123	47.877	52.123	47.877	0	295.000	69.575			
			freq	Z=0	Z=1	Z=0	Z=1	rule	#correct	%correct	p(rule)	p(margin)	

* Rule selected using the independence model.

Input file for SB-search (for demo#4)

dementia05ApEdC.txt: reduce #IVs to 3

```
#Fifth test set of OHSU data 20 June 2012
#Corrected transform of variable I rs12248379
#Corrected transform of APOE
#Updated imputed data
#Deleted var I and dropped records with no var C data
#Added stats on missing. Blank = none missing
```

```
:nominal
ID ,0,0,ID
APOE ,2,1,AP
Gender ,2,0,SX
Education ,3,1,ED
AgeLastExam ,3,0,AG
rs1801133 ,3,0,A
rs3818361 ,4,0,B # missing 3
rs7561528 ,3,1,C
rs744373 ,3,0,D
rs6943822 ,3,0,E
rs4298437 ,3,0,F
rs7012010 ,3,0,G
rs11136000 ,3,0,H
rs10786998 ,4,0,J # missing 9
rs11193130 ,4,0,K # missing 11
rs610932 ,3,0,L
rs3851179 ,3,0,M
rs3764650 ,4,0,N # missing 2
rs3865444 ,4,0,P # missing 9
CaseControl ,2,2,Z
```

```
:no-frequency
```

```
:data
#IndID APOE GENDER EDU ALE A B C D E F G H J K
101 0 0 2 2 1 1 0 1 2 2 1 1 2 0
103 0 0 2 1 0 2 2 0 1 1 1 2 2 0
111 0 1 2 1 2 2 1 1 0 1 1 2 1 1
112 0 0 2 2 2 2 1 1 1 2 1 1 0 2
118 0 1 0 2 2 2 2 0 0 1 1 1 . 1
120 0 1 2 2 2 1 2 1 0 1 1 2 . 1
121 0 0 2 2 2 2 2 1 2 0 0 0 2 0
122 0 0 1 2 2 1 2 1 2 0 0 2 2 0
123 0 0 2 2 2 2 2 0 1 1 0 0 2 0
126 0 0 2 2 2 2 2 0 1 1 2 2 2 0
127 0 1 2 2 2 0 2 1 1 1 1 2 2 0
128 0 0 2 2 2 1 2 1 0 1 1 2 . 2
129 1 1 2 2 2 0 2 1 0 0 1 2 0 2
132 0 0 2 2 2 2 2 1 0 0 1 2 1 1
134 0 1 2 2 2 0 0 2 0 0 1 1 0 2
135 0 0 2 2 2 2 1 1 0 2 1 1 0 2
```

SB search output (demo#4)

ID	MODEL	Level	H	dDF	dLR	Alpha	Inf	%dH (DV)	dAIC	dBIC	Inc.Alpha	Prog.	%C(Data)	%cover
16	IV:Ap0Ed0C0Z:Ap0Ed2C2Z:Ap0Ed0Z:Ap1Z:C2Z:Z	5	4.3897	5	86.7978	0.0000	0.95071864	14.7861	76.7978	56.5491	0.0949	12	69.5755	100.0000
15	IV:Ap0Ed0C0Z:Ap0Ed0Z:Ap0C0Z:Ap1Z:C2Z:Z	5	4.3906	5	86.2578	0.0000	0.94480434	14.6941	76.2578	56.0091	0.1337	12	70.0472	100.0000
14	IV:Ap0Ed2C2Z:Ap1Ed0C0Z:Ap0Ed0Z:Ap1Z:C2Z:Z	5	4.3906	5	86.2410	0.0000	0.94462025	14.6912	76.2410	55.9923	0.1962	13	69.5755	100.0000
13*	IV:Ap0Ed2C2Z:Ap0Ed0Z:Ap1Z:C2Z:Z	4	4.3935	4	84.5686	0.0000	0.92630198	14.4063	76.5686	60.3697	0.0337	9	69.5755	100.0000
12	IV:Ap0Ed0C0Z:Ap0Ed0Z:Ap1Z:C2Z:Z	4	4.3944	4	84.0051	0.0000	0.92012989	14.3103	76.0051	59.8062	0.1044	10	69.5755	100.0000
11	IV:Ap0Ed0Z:Ap0C2Z:Ap1Z:Ed2C2Z:Z	4	4.3951	4	83.6256	0.0000	0.91597252	14.2457	75.6256	59.4266	0.0567	8	69.5755	100.0000
10*	IV:Ap0Ed0Z:Ap1Z:C2Z:Z	3	4.3989	3	81.3650	0.0000	0.89121192	13.8606	75.3650	63.2158	0.0002	7	69.5755	100.0000
9*	IV:Ap0Ed2C2Z:Ap0Ed0Z:Ap1Z:Z	3	4.4011	3	80.0645	0.0000	0.87696704	13.6390	74.0645	61.9153	0.0002	6	69.5755	100.0000
8*	IV:Ap0Ed0Z:Ap0C2Z:Ap1Z:Z	3	4.4013	3	79.9909	0.0000	0.87616134	13.6265	73.9909	61.8417	0.0004	7	69.5755	100.0000
7*	IV:Ap0Ed0Z:Ap1Z:Z	2	4.4216	2	68.0475	0.0000	0.74534205	11.5919	64.0475	55.9480	0.0000	4	69.5755	100.0000
6*	IV:Ap0Ed2C2Z:Ap1Z:Z	2	4.4232	2	67.1062	0.0000	0.73503174	11.4316	63.1062	55.0067	0.0000	2	66.9811	100.0000
5*	IV:Ap1Z:Ed0Z:Z	2	4.4243	2	66.4783	0.0000	0.72815386	11.3246	62.4783	54.3788	0.0001	4	69.5755	100.0000
4*	IV:Ap1Z:Z	1	4.4504	1	51.0910	0.0000	0.55961325	8.7034	49.0910	45.0413	0.0000	1	66.9811	100.0000
3*	IV:Ap0C2Z:Z	1	4.4727	1	37.9865	0.0000	0.41607561	6.4710	35.9865	31.9367	0.0000	1	62.2642	100.0000
2*	IV:Ap0Ed2C2Z:Z	1	4.4770	1	35.4733	0.0000	0.38854880	6.0429	33.4733	29.4236	0.0000	1	58.7264	100.0000
1*	IV:Z	0	4.5374	0	0.0000	1.0000	0.00000000	0.0000	0.0000	0.0000	0.0000	0	52.1226	100.0000

ID	MODEL	Level	H	dDF	dLR	Alpha	Inf	%dH (DV)	dAIC	dBIC	Inc.Alpha	Prog.	%C(Data)	%cover
----	-------	-------	---	-----	-----	-------	-----	----------	------	------	-----------	-------	----------	--------

Best Model(s) by dBIC:

10*	IV:Ap0Ed0Z:Ap1Z:C2Z:Z	3	4.3989	3	81.3650	0.0000	0.89121192	13.8606	75.3650	63.2158	0.0002	7	69.5755	100.0000
-----	-----------------------	---	--------	---	---------	--------	------------	---------	---------	---------	--------	---	---------	----------

Best Model(s) by dAIC:

16	IV:Ap0Ed0C0Z:Ap0Ed2C2Z:Ap0Ed0Z:Ap1Z:C2Z:Z	5	4.3897	5	86.7978	0.0000	0.95071864	14.7861	76.7978	56.5491	0.0949	12	69.5755	100.0000
----	---	---	--------	---	---------	--------	------------	---------	---------	---------	--------	----	---------	----------

Best Model(s) by Information, with all Inc. Alpha < 0.05:

13*	IV:Ap0Ed2C2Z:Ap0Ed0Z:Ap1Z:C2Z:Z	4	4.3935	4	84.5686	0.0000	0.92630198	14.4063	76.5686	60.3697	0.0337	9	69.5755	100.0000
-----	---------------------------------	---	--------	---	---------	--------	------------	---------	---------	---------	--------	---	---------	----------