Bayesian Knowledge Bases (BKBs)

- What is a BKB
- What is an Inference
- Probabilistic Reasoning with BKBs
- Uniquely Representing a BKB as a matrix
- Different types of BKBs

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Richard S. Detsch

Thayer School of Engineering

October 2006

Richard.S.Detsch@dartmouth.edu

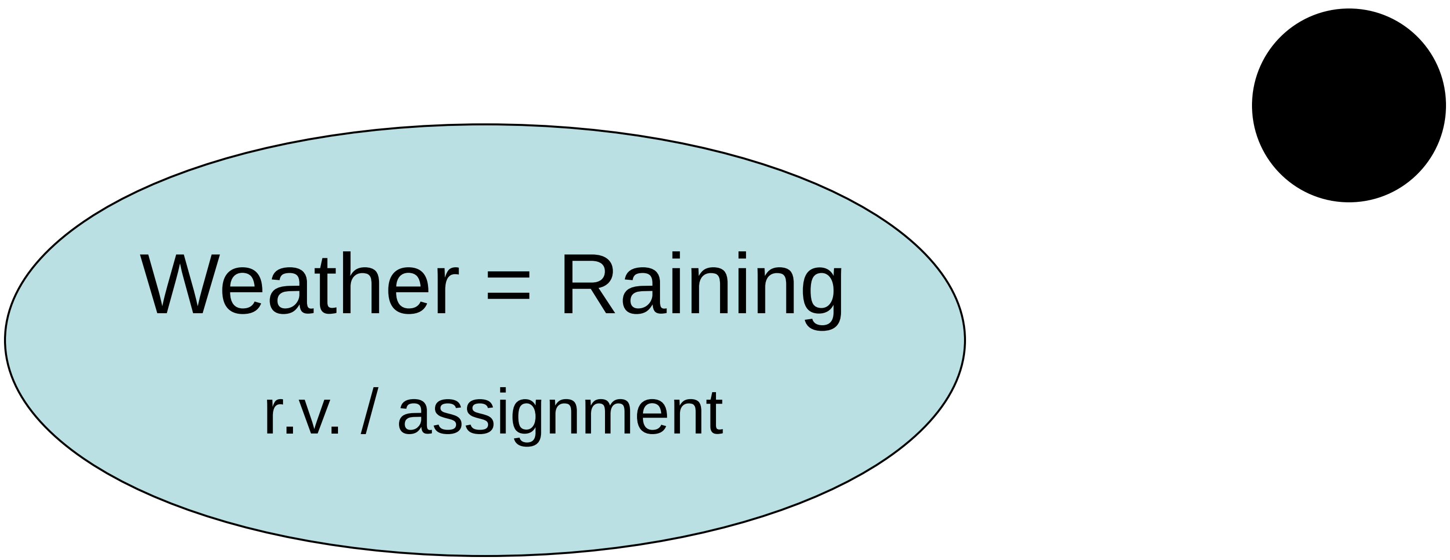
Slides on paper titled: Reasoning with BKBs – Algorithms and Complexity by Rosen, Santos, Shimony

Why Construct a BKB

- Similar to why you would want to construct any knowledge network: *to find the most probable complete state (belief revision) & associated marginal probabilities (belief updating) given the evidence.*
- Belief Updating schemes do not seem to work well with BKBs ☹

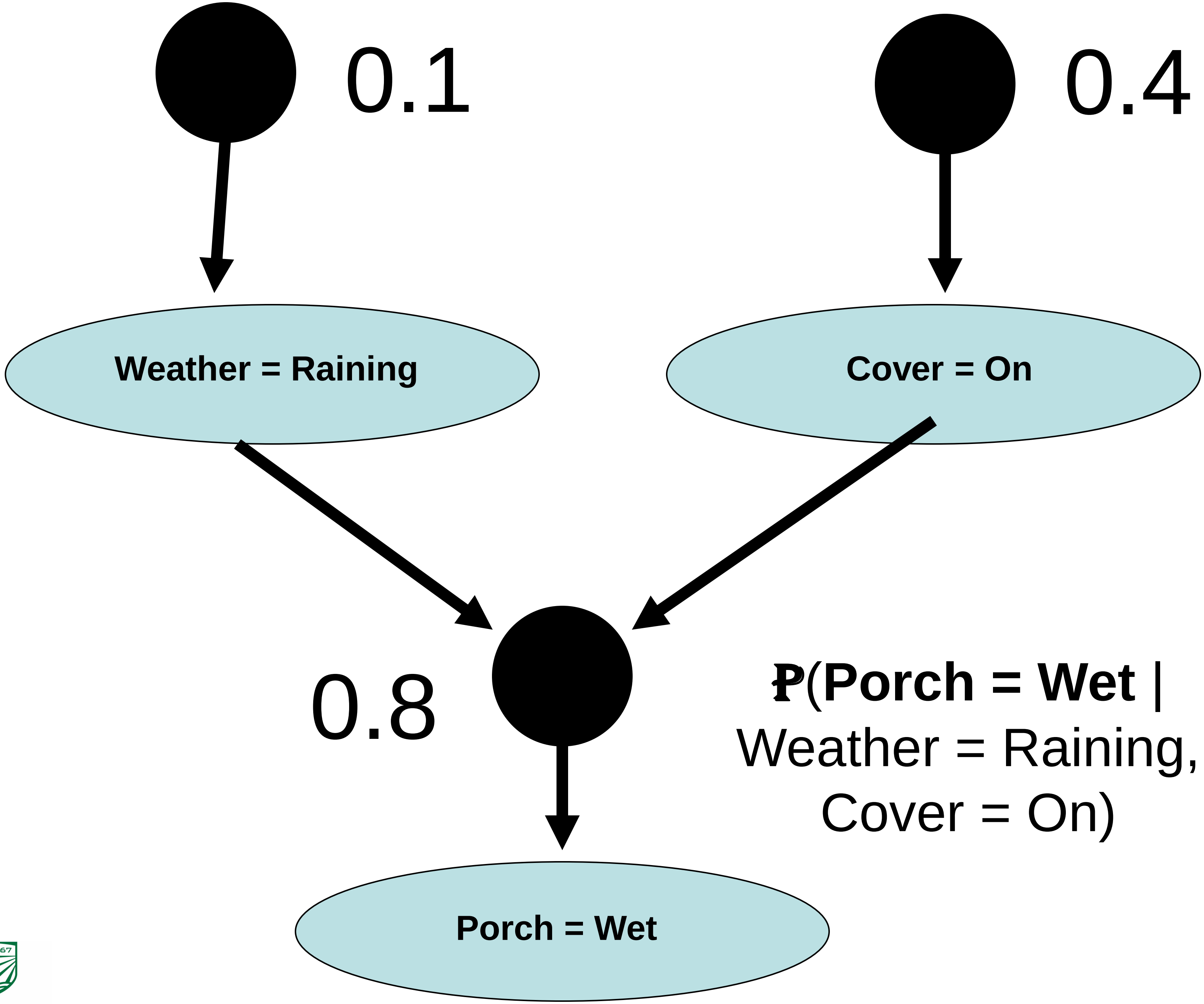
What is a BKB

S-node “Conditional Probability Node”



I-Node “Instantiation Node”

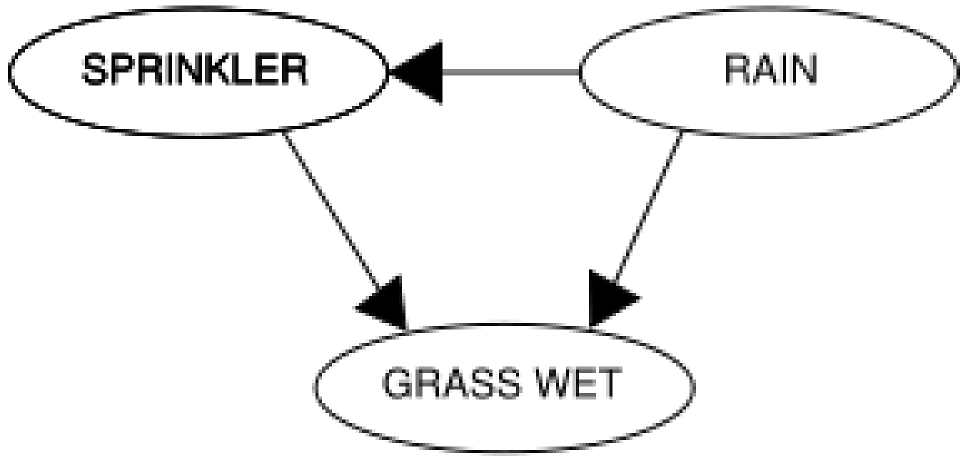
THAYER SCHOOL OF ENGINEERING AT DARTMOUTH



$\mathbb{P}$(Porch = Wet | Weather = Raining, Cover = On)

Bayesian Network

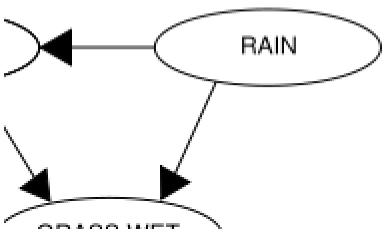
	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



RAIN	
T	F
0.2	0.8

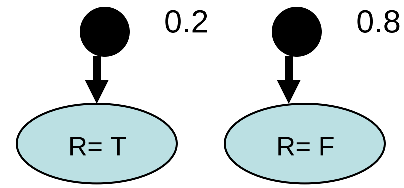
		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

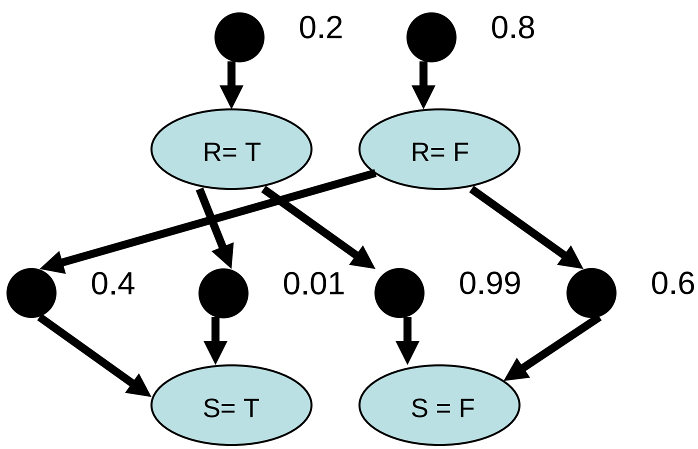
BN to BKB



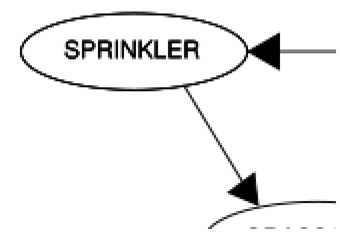
RAIN	
T	F
0.2	0.8

=



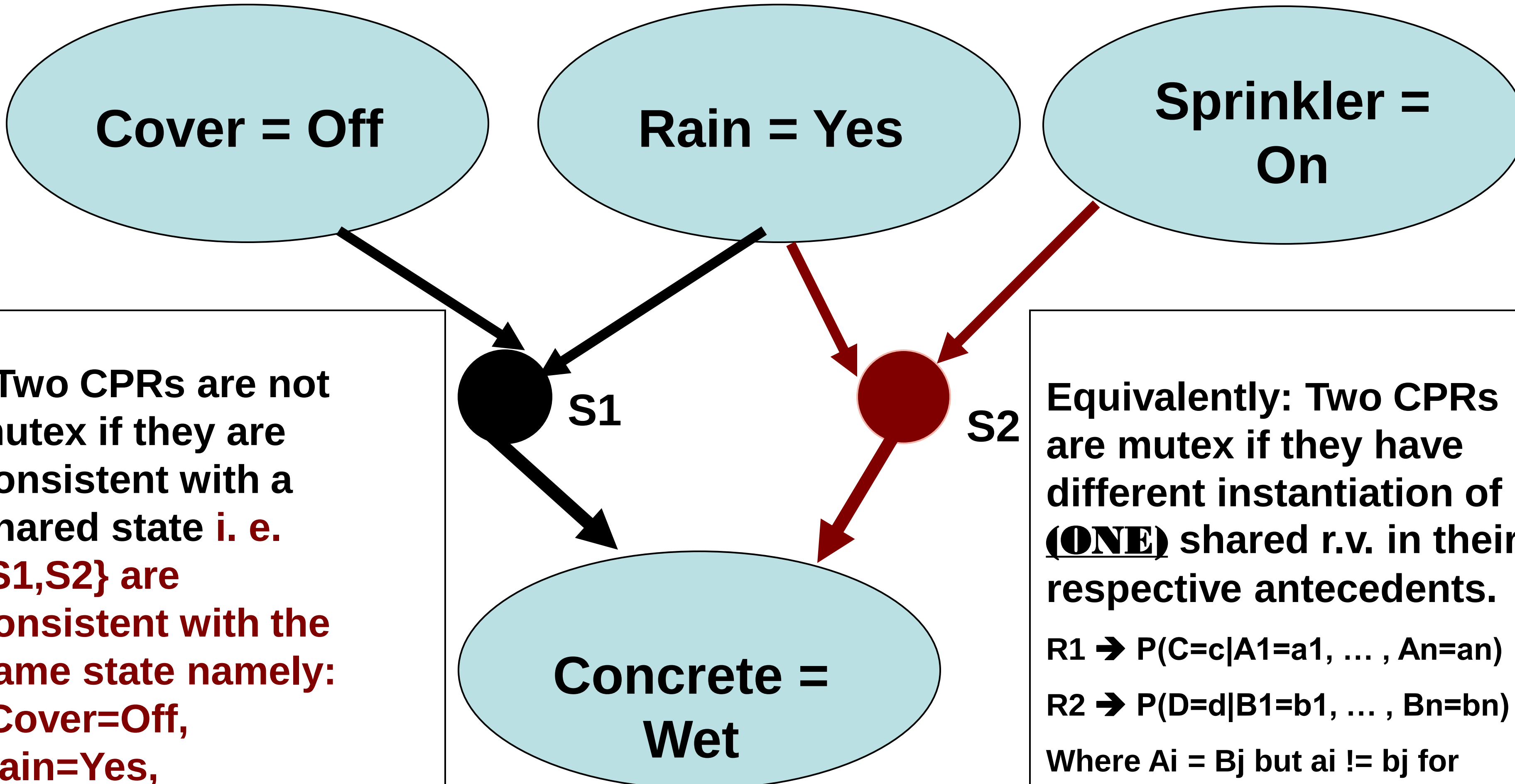


	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



For each (***variable***, value) pair construct an ***I-node r***. For each CPT entry construct a rule with the antecedent being the I-node corresponding to the state of the predecessor *(could be null)* of ***variable***, and consequent being the ***I-node r***

? Mutual Exclusion ?



Two CPRs are not mutex if they are consistent with a shared state i. e. {S1,S2} are consistent with the same state namely: {Cover=Off, Rain=Yes, Sprinkler=On, Concrete=Wet}

Equivalently: Two CPRs are mutex if they have different instantiation of <u>**(ONE)**</u> shared r.v. in their respective antecedents.

R1 ➔ P(C=c|A1=a1, … , An=an)

R2 ➔ P(D=d|B1=b1, … , Bn=bn)

Where Ai = Bj but ai != bj for some i and j

{S1,S2} doesn't have a shared r.v.

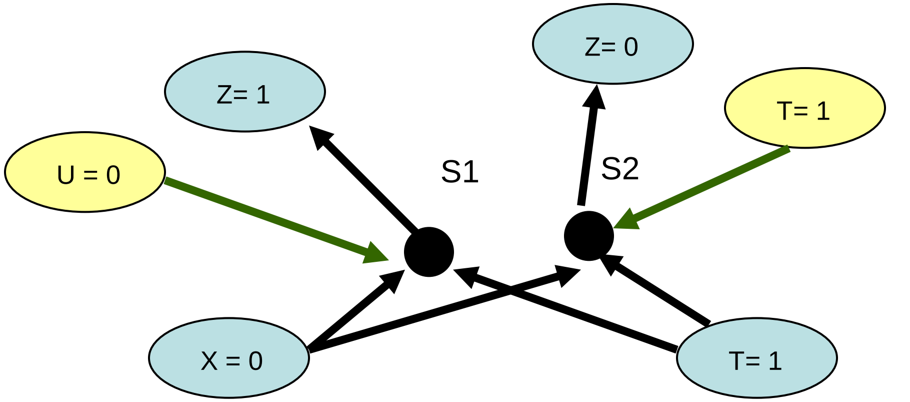
Consequent Bound

P(C = c | A1, … , Am)

P(D = d | B1, … , Bn)

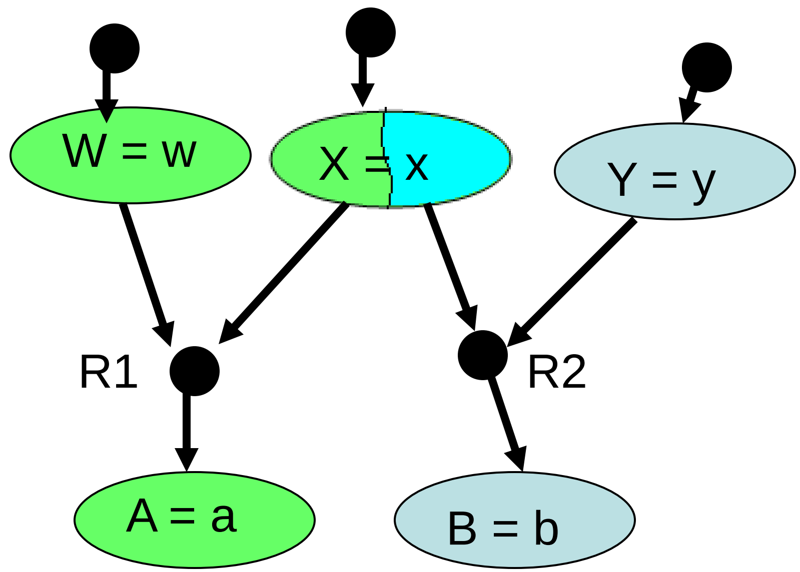
Where ai = aj whenever Ai = Bj and C = D but c != d

Note: Mutually Consequent Bound CPRs are not mutually exclusive



The CPRs {S1,S2} are consequent bound because they have discrete instantiations of the same r.v. in their consequence and they're antecedent intersection is non-null. Therefore **{S1,S2} are opposing rules** to apply when both antecedents are satisfied since they result in different consequents. **Note: the yellow is to indicate that Consequent Bound does not mean Consequent Variant**

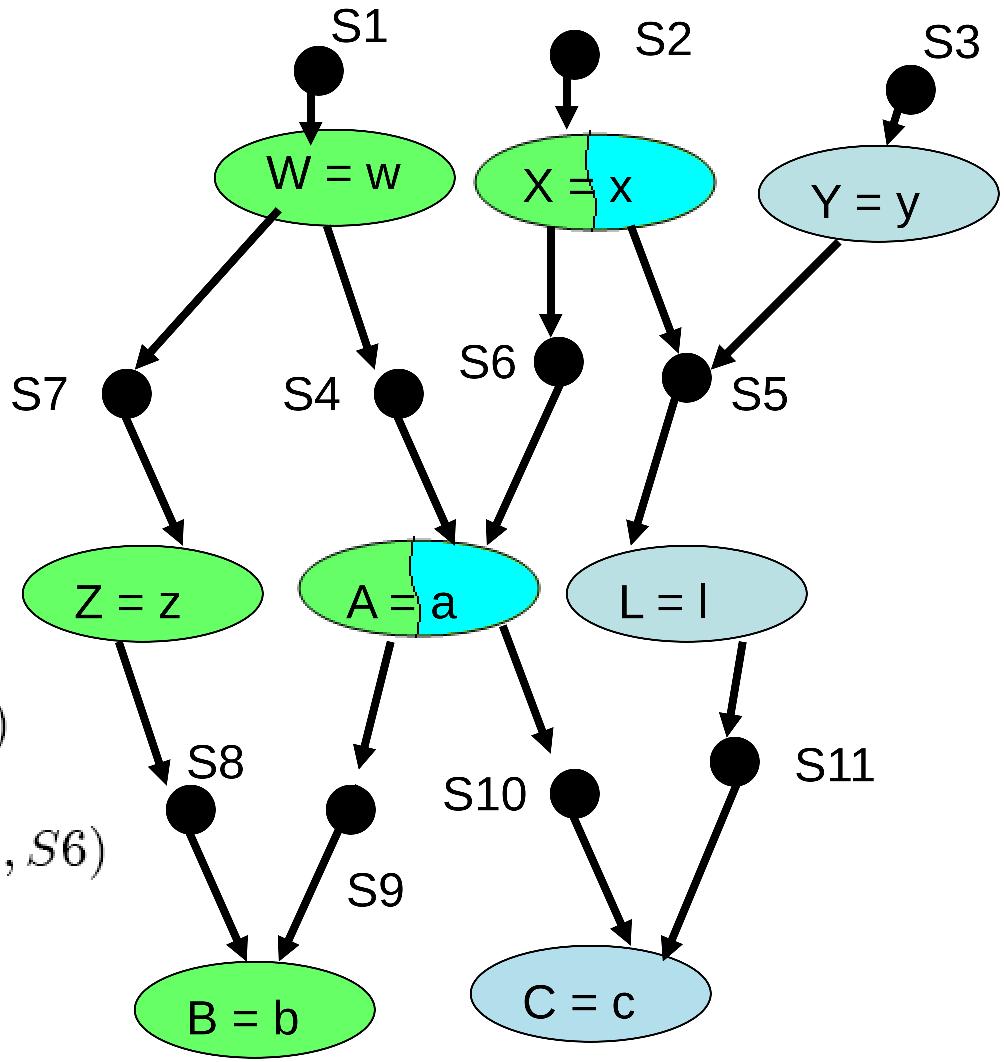
Compatible CPRs



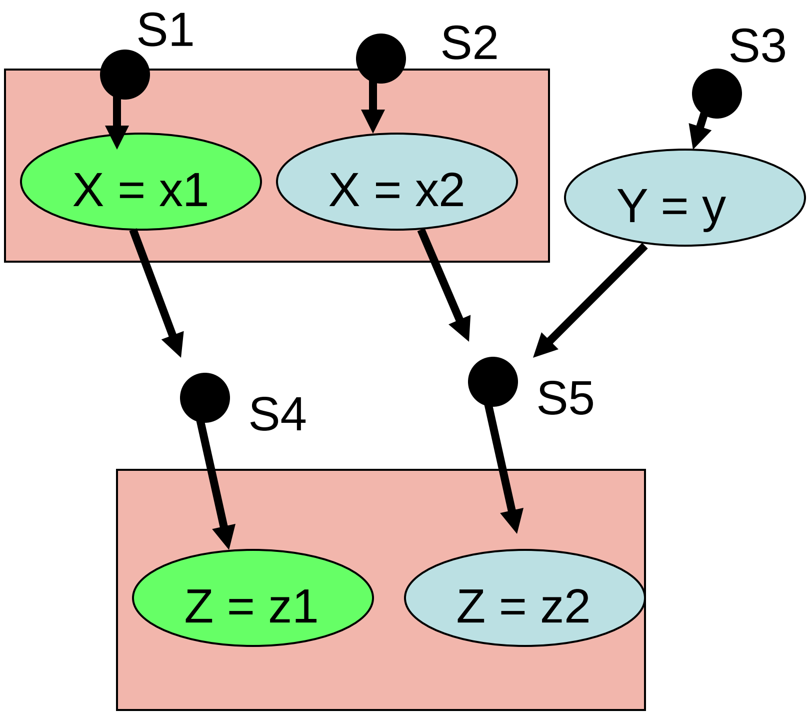
Two CPRs are compatible if either they do not share a r.v. in their respective antecedent or if they do share one or more r.v.s. in their respective antecedents, those r.v.s. have the same instantiation.

Compatible Intersection

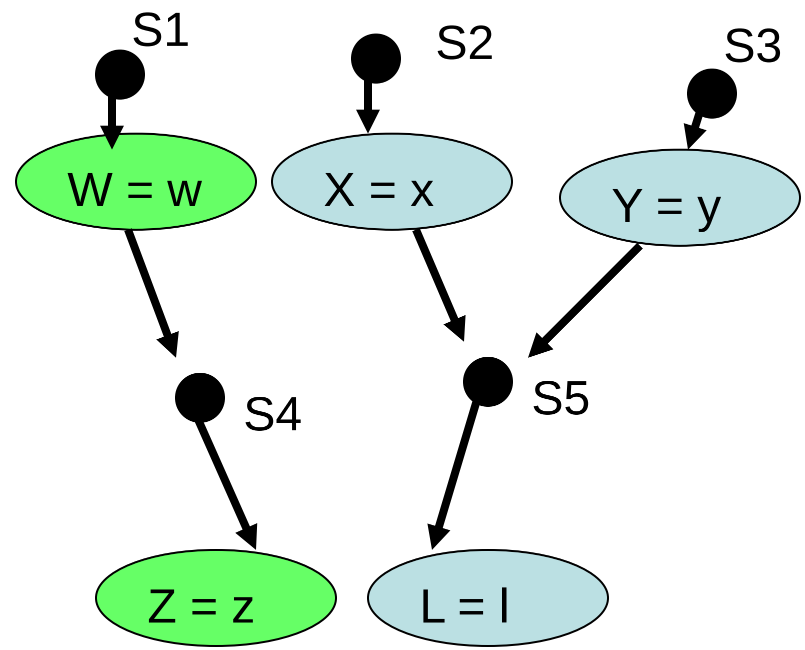
Given two compatible inference Γ_1 & Γ_2
Where $\Gamma_1 = \{S_1, I_1, E_1\}$
and $\Gamma_2 = \{S_2, I_2, E_2\}$
Then their intersection Γ is an inference
Where
$\Gamma = (\{I_1 \cap I_2\} \cup \{S_1 \cap S_2\}, E_1 \cap E_2)$
Here
$\Gamma = (\{X = x, A = a\} \cup \{S2, S6\}, S2, S6)$



Compatible Inferences



Not Compatable



Compatable

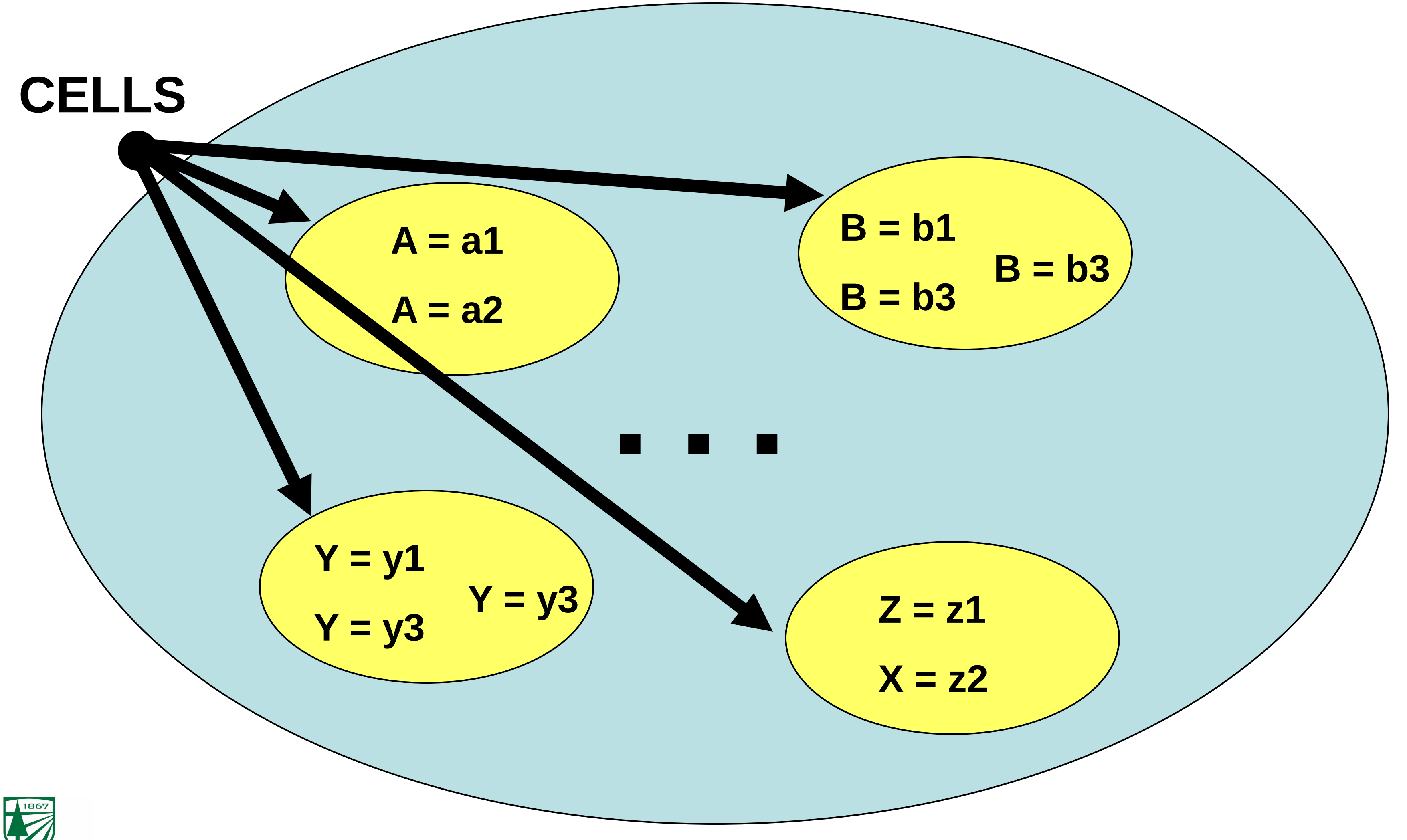
Two inferences I_1 and I_2 are compatible if $\forall\ (i_1, i_2) \in \{I_1 \otimes I_2\}\ (i_1, i_2)$ is compatible

Formal Definition of a BKB $\mathbb{B}$

- For any two distinct CPRs S1 and S2 in $\mathbb{B}$ are either mutually exclusive or have different consequence.
- For any subset of **mutually consequent-bound CPRs** $\mathbb{R}$, the weight of $\mathbb{R}$, $\mathbb{W}(\mathbb{R})$ is less than unity.

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

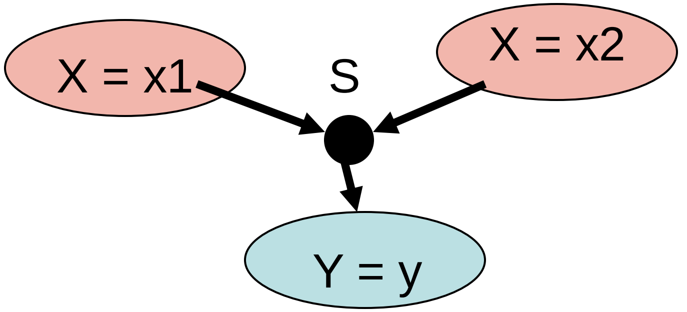
Partition Rule



THAYER SCHOOL OF
ENGINEERING
AT DARTMOUTH

A BKB that obeys a partition ***rule*** in the I-node space if it is confined by two restrictions

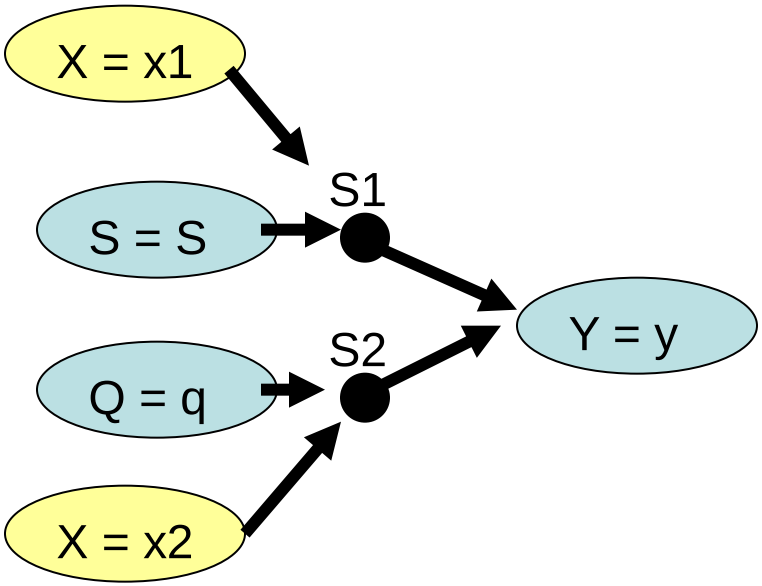
- **One:** No more than one instantiation per r.v. can be contained in any of its S-node's immediate predecessor set



The S-node S has two instantiations of r.v. X represented in its immediate predecessor set

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

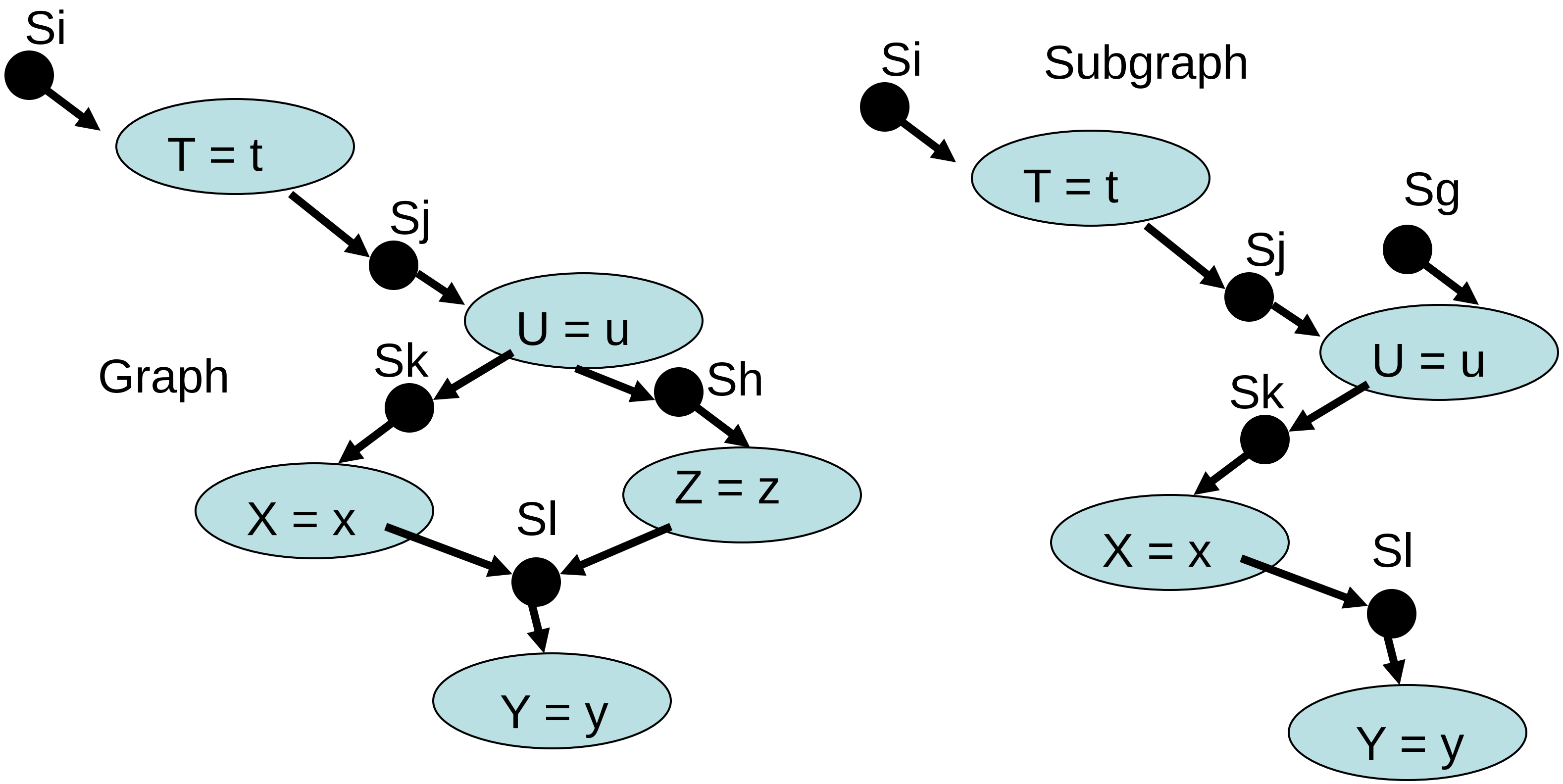
- **Two:** An I-node with two distinct S-nodes in its immediate predecessor set will have two distinct r.v. instantiations for a single r.v.



Note: a set of CPRs that obey a partition rule is said to be mutually exclusive w.r.t. the partition rule; for it doesn't allow events that overlap in probability space for the same I-node

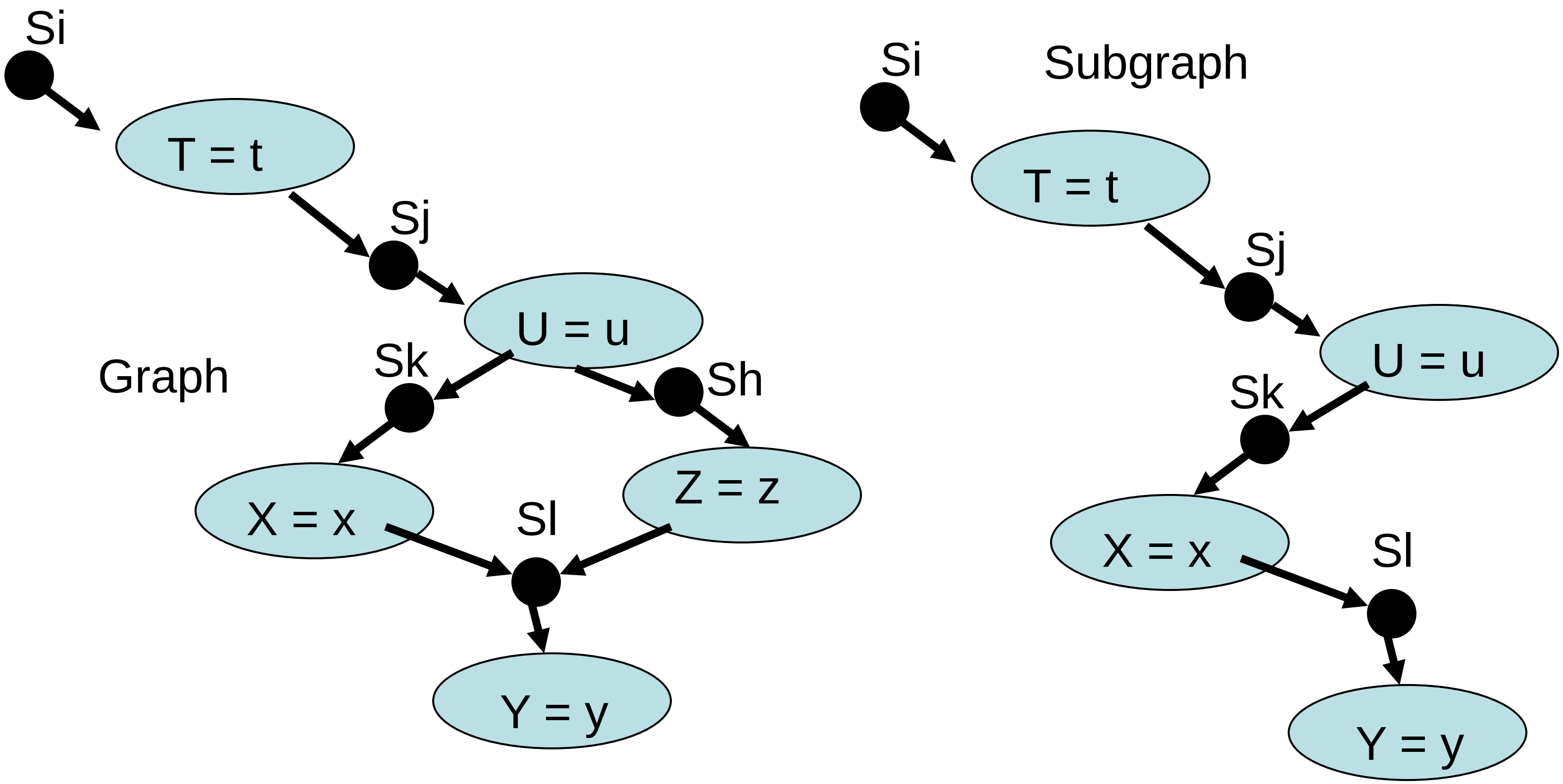
The S-nodes S1 and S2 have r.v. instantiations x1 and x2 represented in their respective immediate predecessor set and therefore are not mutex w.r.t. the partition rule

An S-node within a subgraph is ***well-founded*** if it keeps its original immediate predecessor set.



Here Sk is well-founded in the subgraph but Sl is not

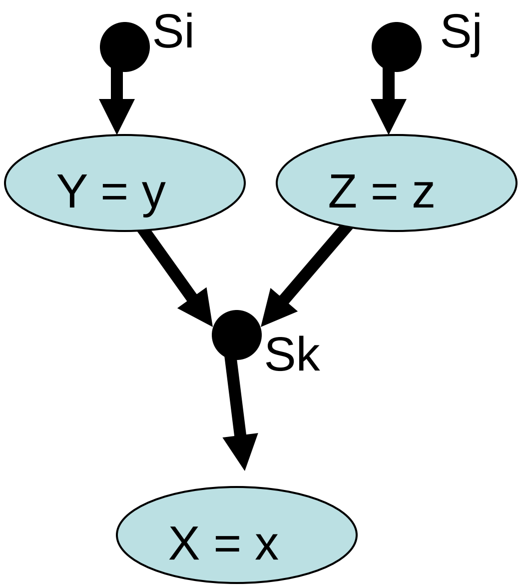
An S-node within a subgraph is ***well-defined*** if it keeps its original immediate descendent set.



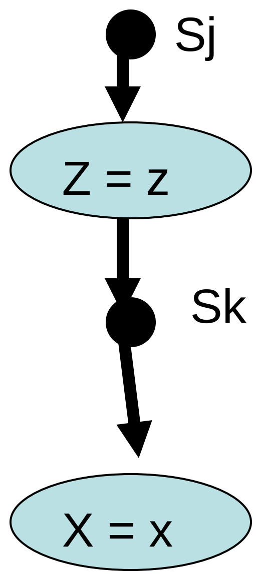
Note: all S-nodes within a subgraph are well-defined because they have to have exactly one I-node descendent to even exist

An I-node is ***well-supported*** if it is preceded by an S-node

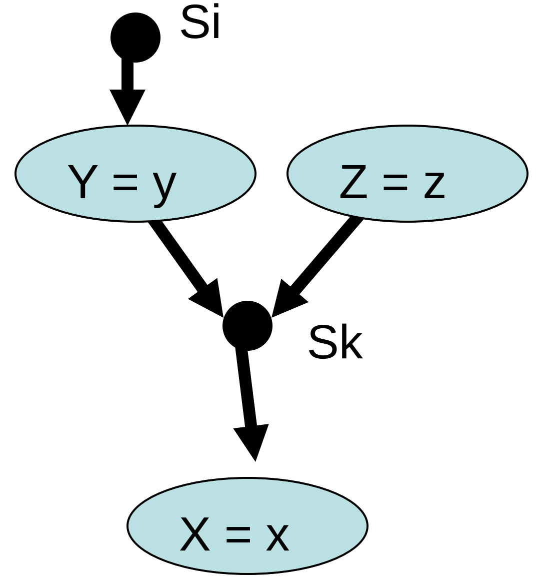
Note: outside of this slide assume all I-nodes not pictured with a preceding S-node to have one implicitly.



Good



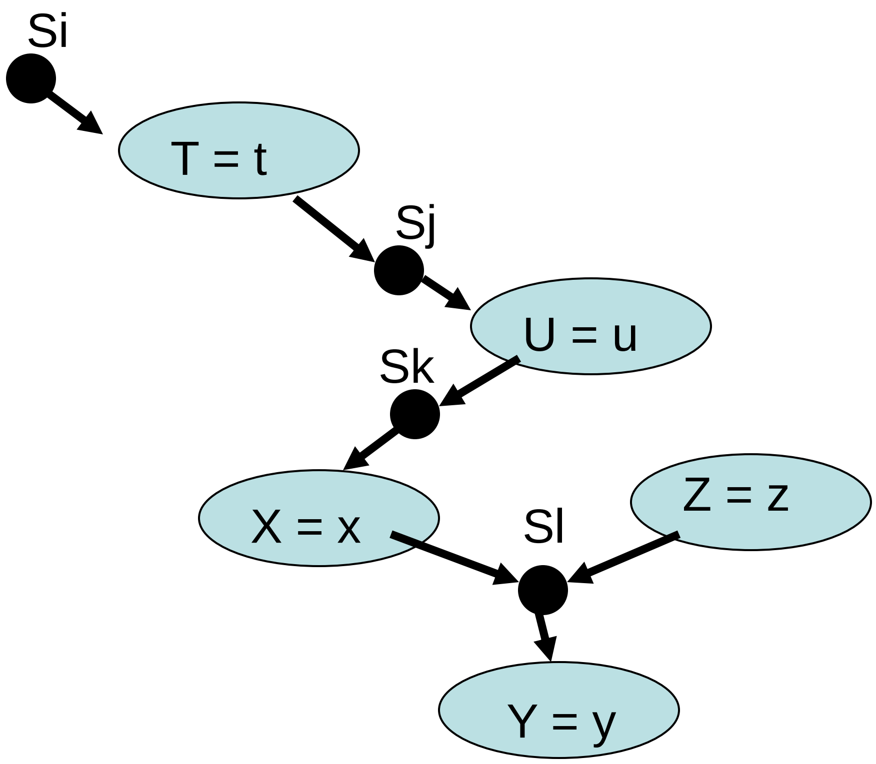
Good



Bad ➔ Z = z not preceded by an S-node

States are subgraphs of a BKB that do not contain more that one r.v. instantiation

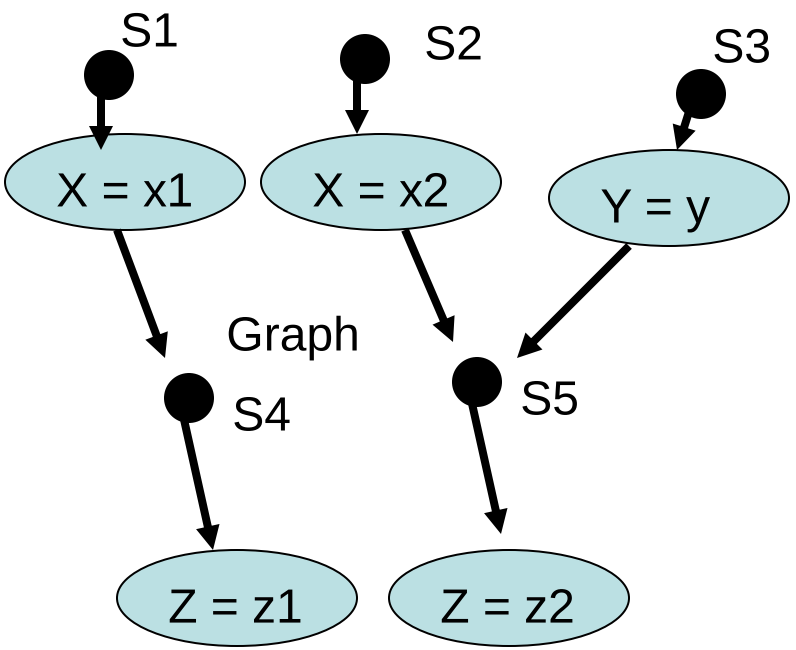
- Every r.v. has two or more possible instantiations
- A ***state*** is a set of I-nodes containing one or less instantiations from each r.v.
- A ***state*** is complete for a set of r.v.s. if it contains exactly one instantiation from each r.v.



Above ***state*** is complete on the set of r.v.s. {T,U,X,Y,Z}

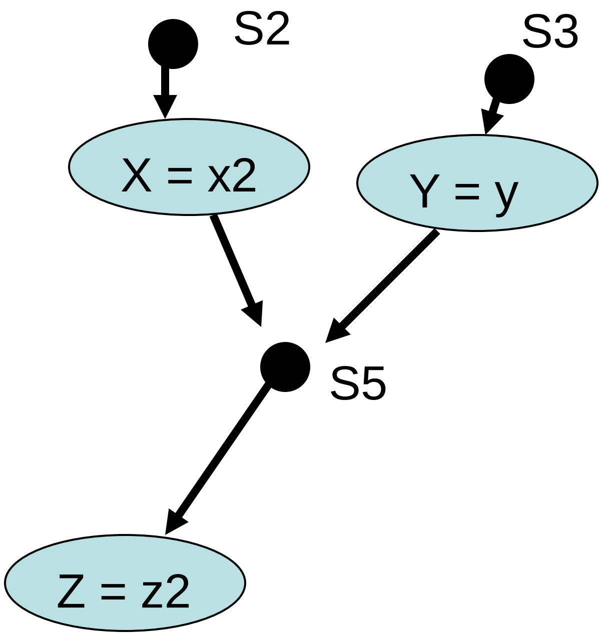
THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

A subgraph is an ***inference*** over a graph if every S-node is well-founded/defined, every I-node is well supported, acyclic and is a ***state***



Because of the partition rule, an I-node with more than one predecessor S-node is not allowed!

A node is grounded if it exist in an inference



Subgraph

This subgraph is ***complete inference*** for it is also a ***complete state*** on r.v.s. {X, Y, Z}, also notice that the graph is not an ***inference*** for it is not a ***state***

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Inferencial Equivalence

Given any two inferences

$$\Gamma_1 = \{S_1, I_1, E_1\}, \Gamma 2 = \{S_1, I_1, E_1\}$$

If either of these are true:

$$S_1 \subseteq S_2 \vee I_1 \subseteq I_2 \vee E_1 \subseteq E_2$$

Then $\Gamma_1 \subseteq \Gamma_2$

When is an inference relevant to a state?

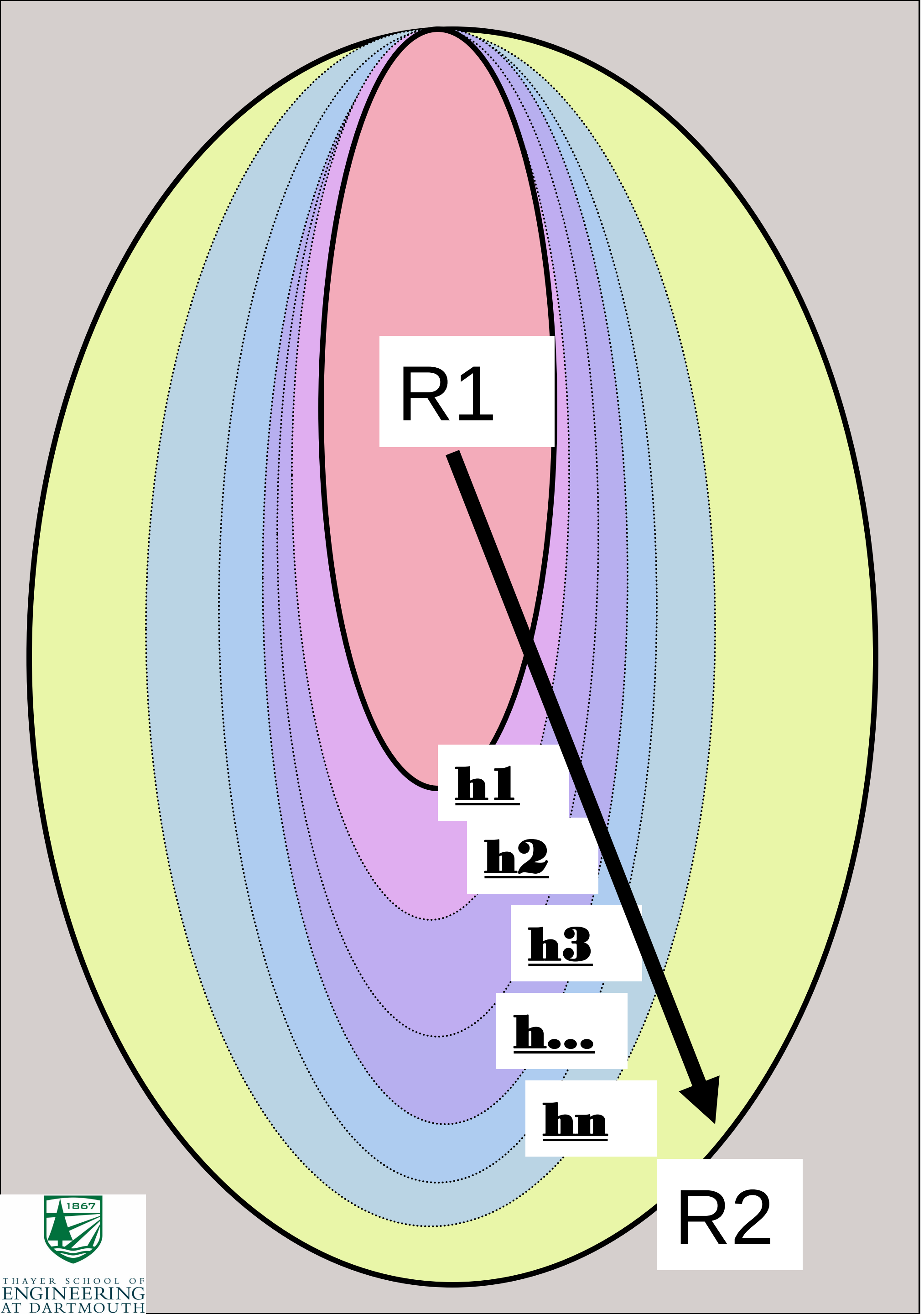
- An inference is relevant to a state when it is a subgraph of that state.
- Multiple inferences can be relevant to a state but their exist one unique maximally relevant inference, however multiple states can share the same maximally relevant inference
- The **composite state of an inference** is the set of complete states to which the inference is relevant. **Dominate composite state** of an inference is the set of complete states the inference is the maximal relevant inference to.

BKB Locally Complete?

$\mathbb{R}$: Y1= y1 Y2= y2 • • • Yk= yk ➔ X = xj

When you have a rule $\mathbb{R}$ s.t. given any complete state $\mathbb{S}$ (on ant($\mathbb{R}$)) with non-zero probability then ant($\mathbb{R}$) {Y1=y1, … ,Yk=yk} extended by X=xl will have probability less then $\mathbb{S}$ there exist an inference $\mathbb{I}$ in the BKB relevant to $\mathbb{S}$ that contains ant($\mathbb{R}$) + X=xl

$\mathbb{I}$: Y1= y1 Y2= y2 • • • Yk= yk + X = xl



THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

For any two inferences **R1** and **R2** where **R1** is a proper subset of **R2** there exist a sequence of inferences

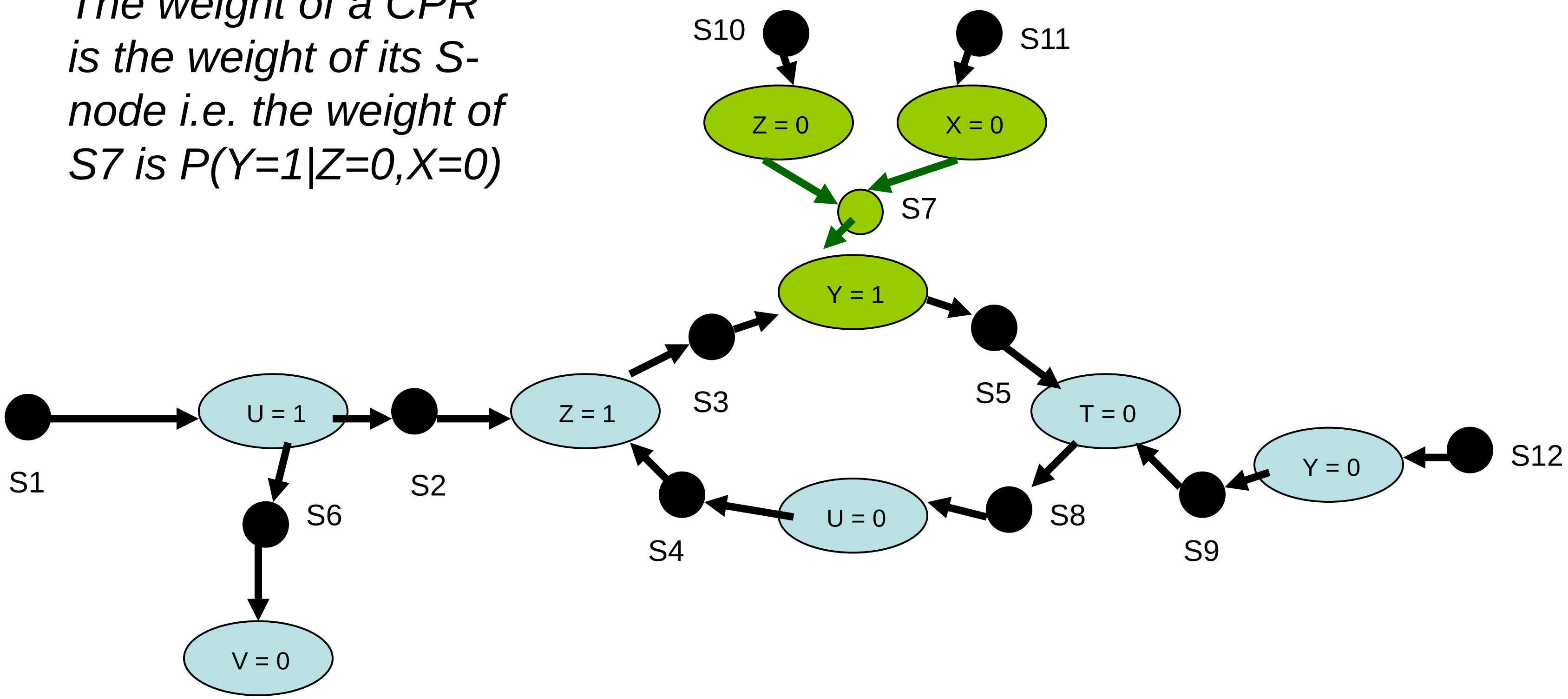
{h1, h2, h3, … , hn} called the **immediate parent sequences** where:

$|\mathbb{I}_{hi+1} - \mathbb{I}_{hi}| = |\mathbb{S}_{hi+1} = \mathbb{S}_{hi}| = 1$

that formally defines in a discrete incremental fashion the canonical inference space between the two inferences described above.

Conditional Probability Rule (CPR)

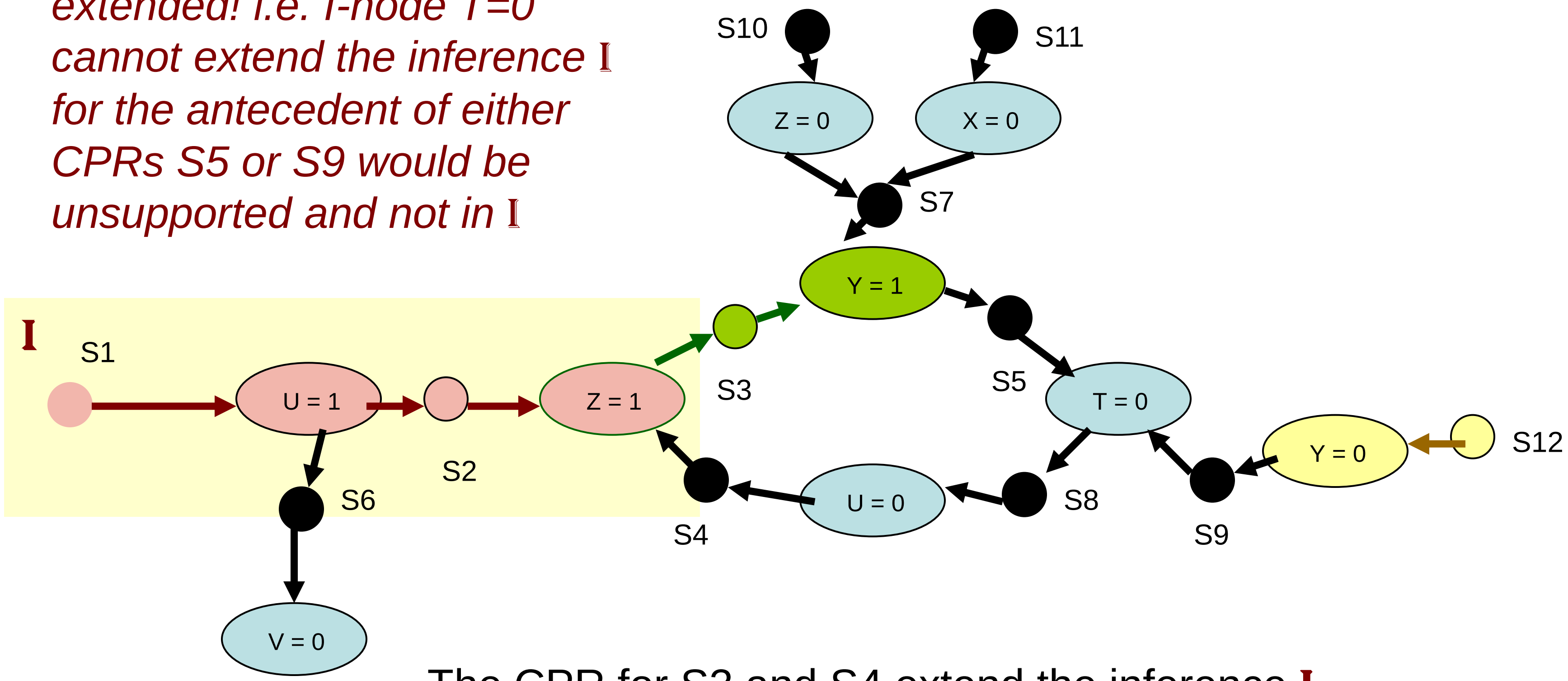
The weight of a CPR is the weight of its S-node i.e. the weight of S7 is P(Y=1|Z=0,X=0)



A CPR is an S-node, its incident edges and immediate neighbors; highlighted in green above is the CPR for S-node S7

A CPR is called an ***extender*** of an inference if it is not a member of the inference but when include creates a new inference.

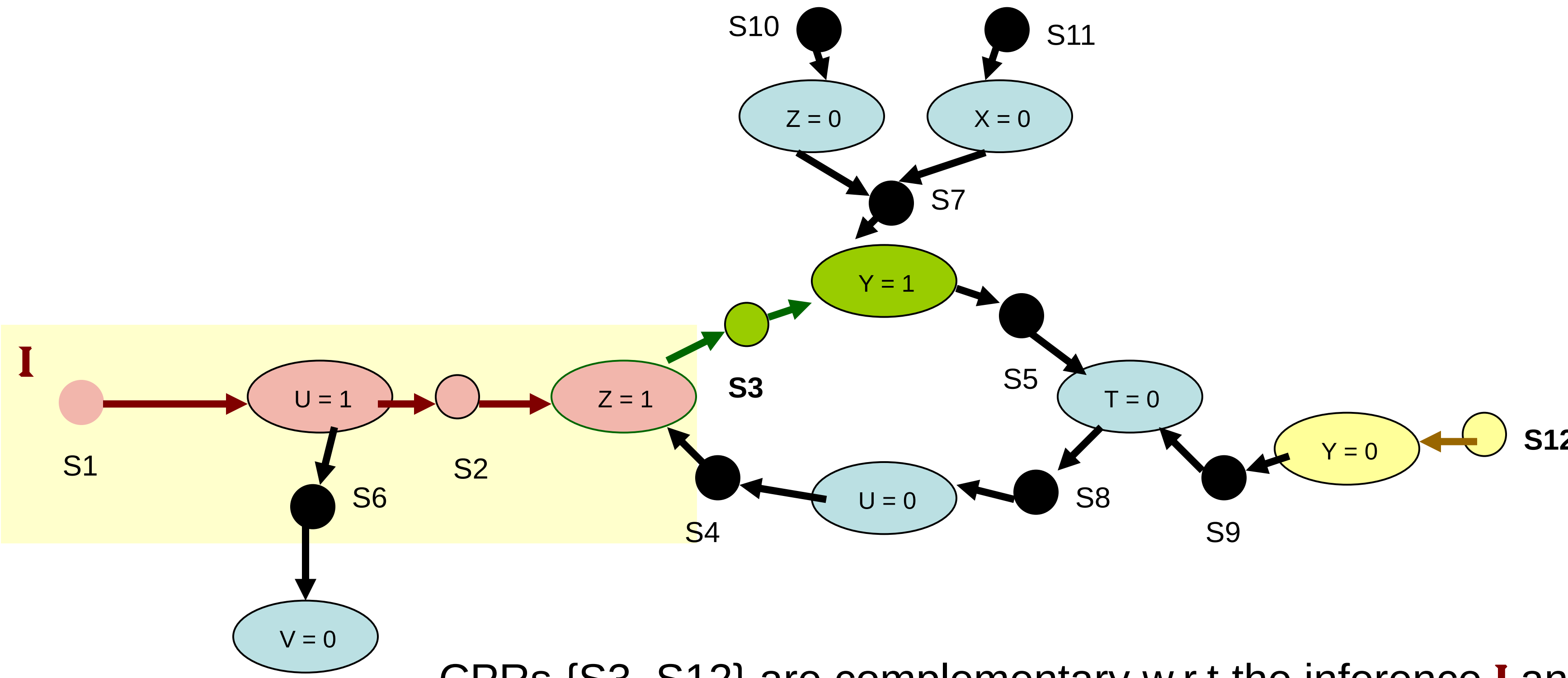
Not all r.v.s that are outside the inference can be extended! i.e. I-node T=0 cannot extend the inference **I** *for the antecedent of either CPRs S5 or S9 would be unsupported and not in* **I**



The CPR for S3 and S4 extend the inference **I**.

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

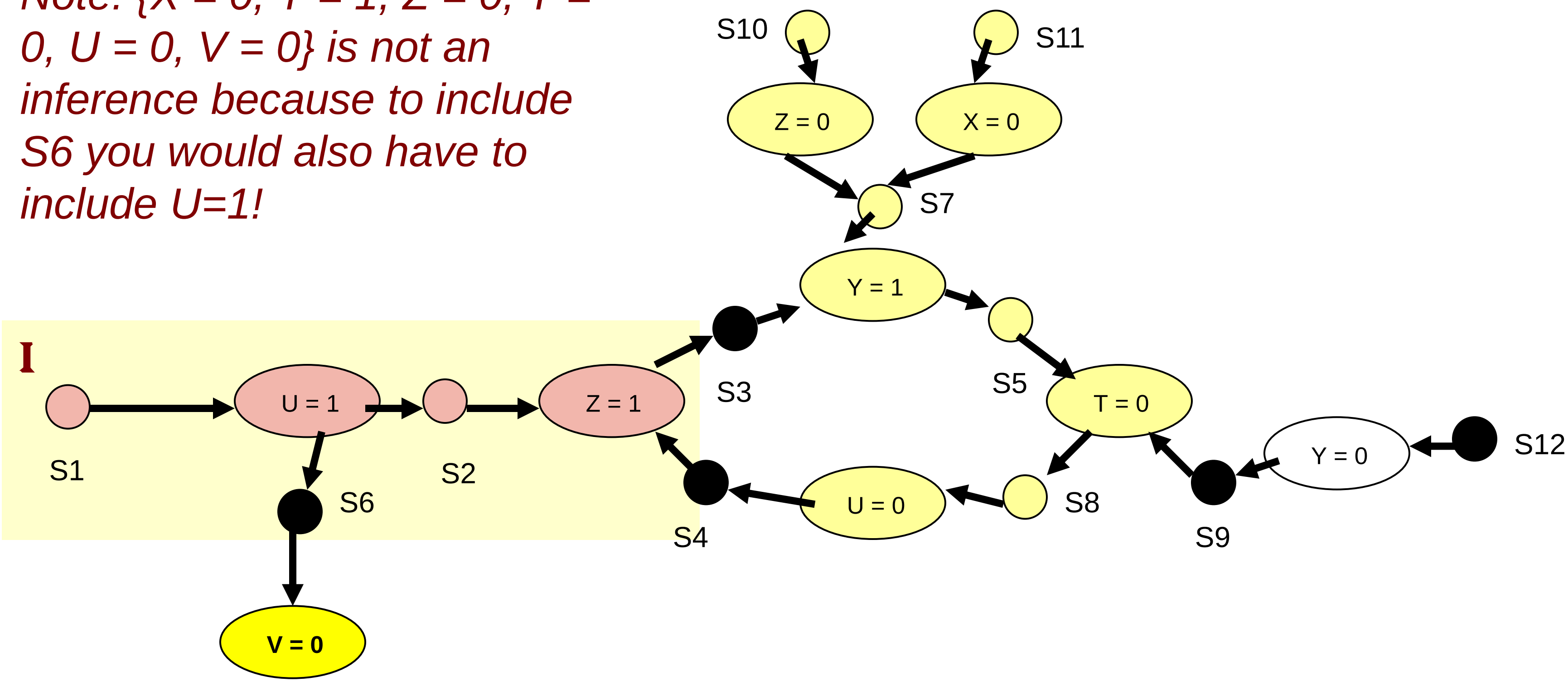
A set of CPRs is **complementary** w.r.t. an inference and a r.v. if each extents the inference by including a unique instantiation of the r.v.



CPRs {S3, S12} are complementary w.r.t the inference **I** and the r.v. Y additionally if Y=0 and Y=1 were the only instantiations for Y then {S3, S12} is the unique maximal complementary set of CPRs.

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Note: {X = 0, Y = 1, Z = 0, T = 0, U = 0, V = 0} is not an inference because to include S6 you would also have to include U=1!



The inference **I** is relevant to the following states: {X = 0, Z = 1, U = 1};

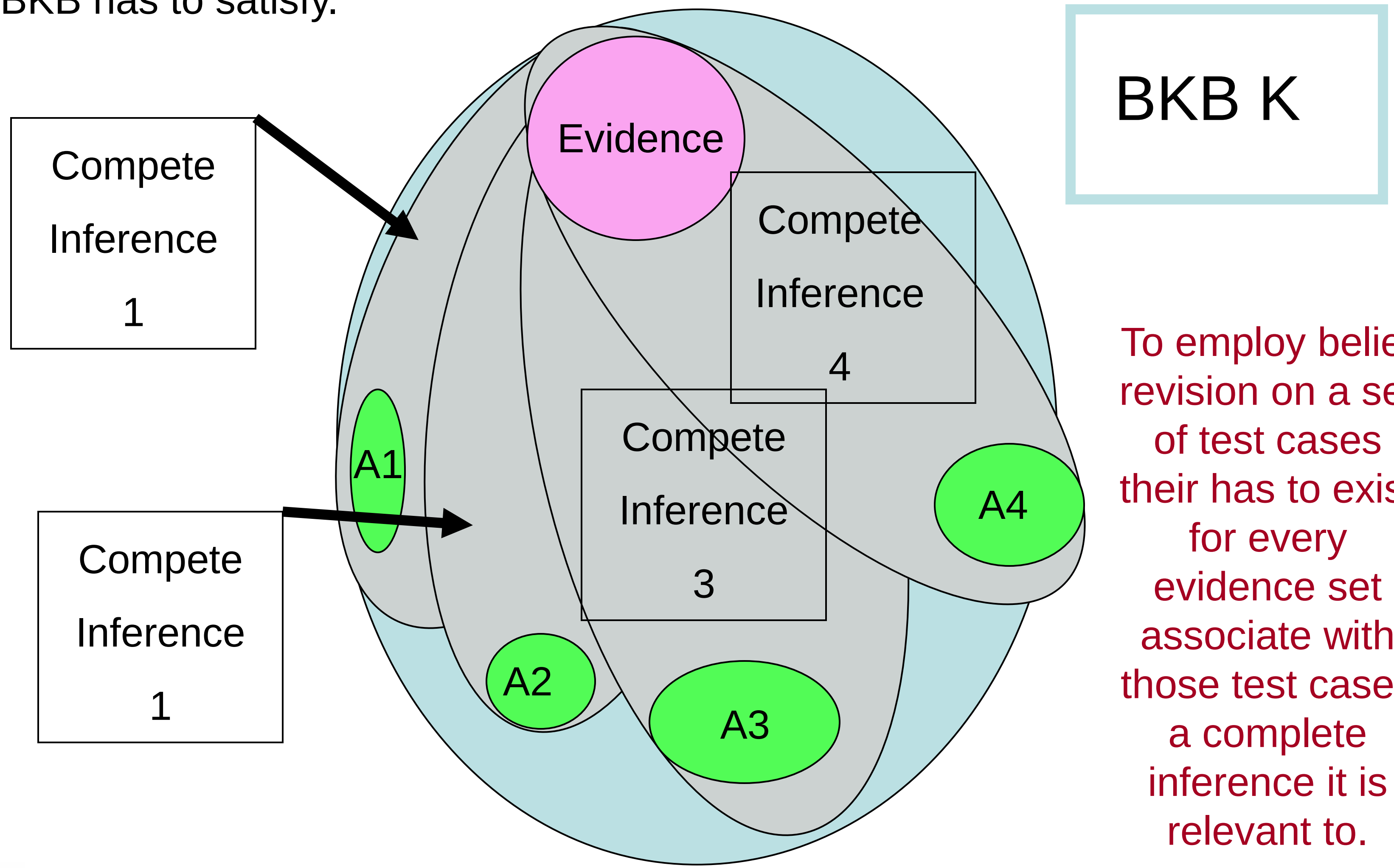
{X = 0, Y = 2, Z = 1, T = 0, U = 1}.

The inference in light yellow is the maximally relevant to the state

{X = 0, Y = 1, Z = 0, T = 0, U = 0, V = 0}. "*which is a member of the dominate composite state of the inference*"

The user defines what test cases and associated evidence that the BKB has to satisfy.

Test Case

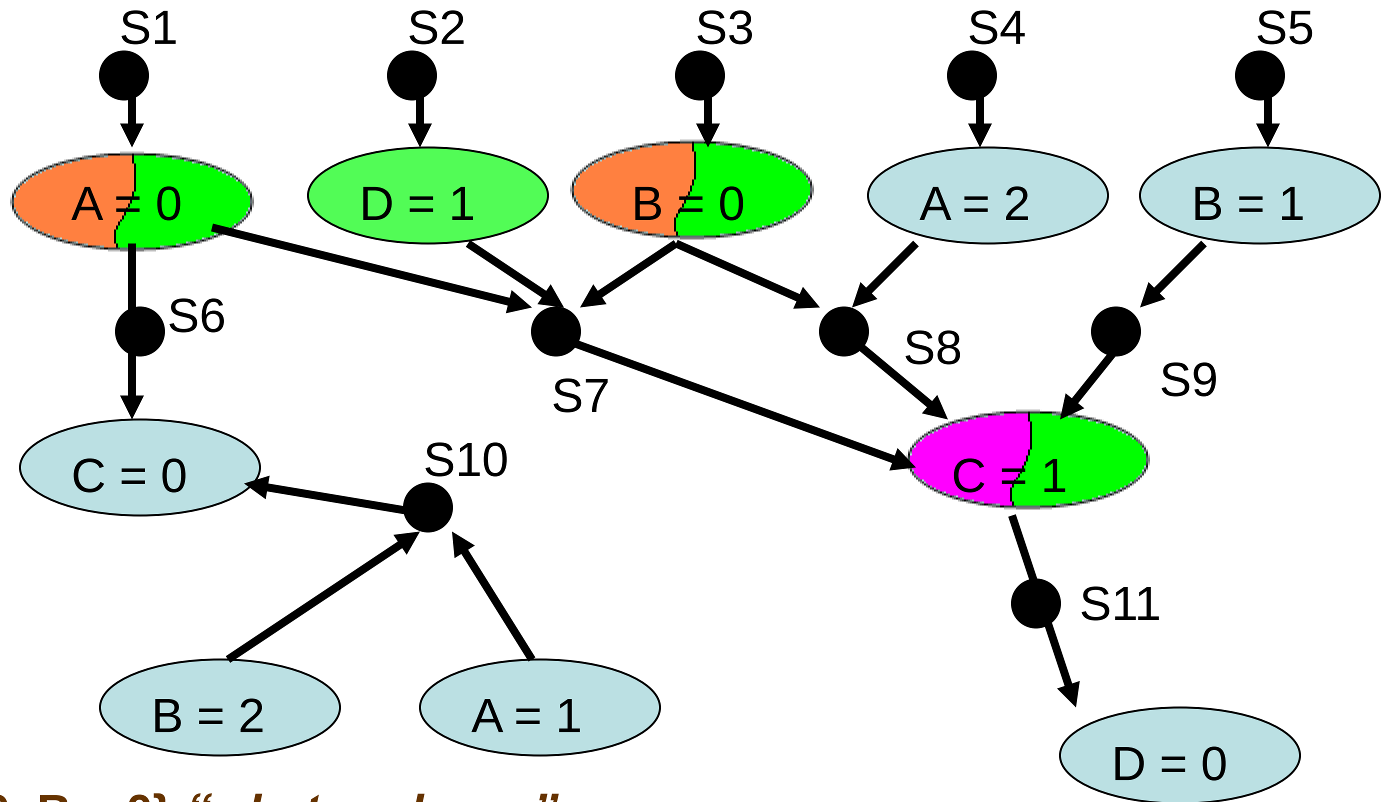


To employ belief revision on a set of test cases their has to exist for every evidence set associate with those test cases a complete inference it is relevant to.

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

BKB User Mechanics

The evidence is the customers query which through belief revision produces the most probable compete state containing the evidence, the answer are the r.v. values in the complete state that have particular interest to the user.



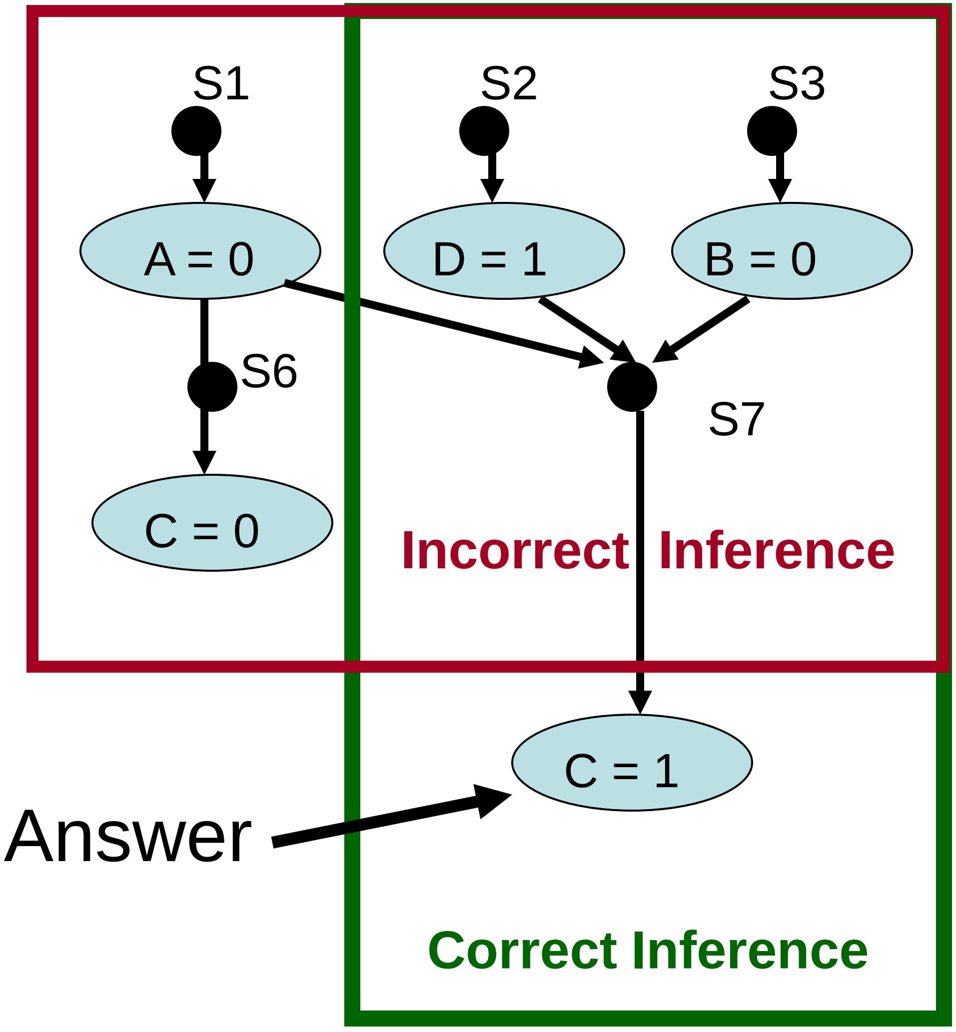
Evidence E ➔ {A = 0, B = 0} *"what we know"*

MPE *"most probable complete state"* {A = 0, B = 0, D = 1, C = 1}

The random variable the customer cares about C ➔ Answer = {C = 1} if C=1 is the expected answer and P(C = 1|E) > P(C = * | E) where * is all other possible values for r.v. C then the BKB satisfies the Evidence in respect to the expected answer.

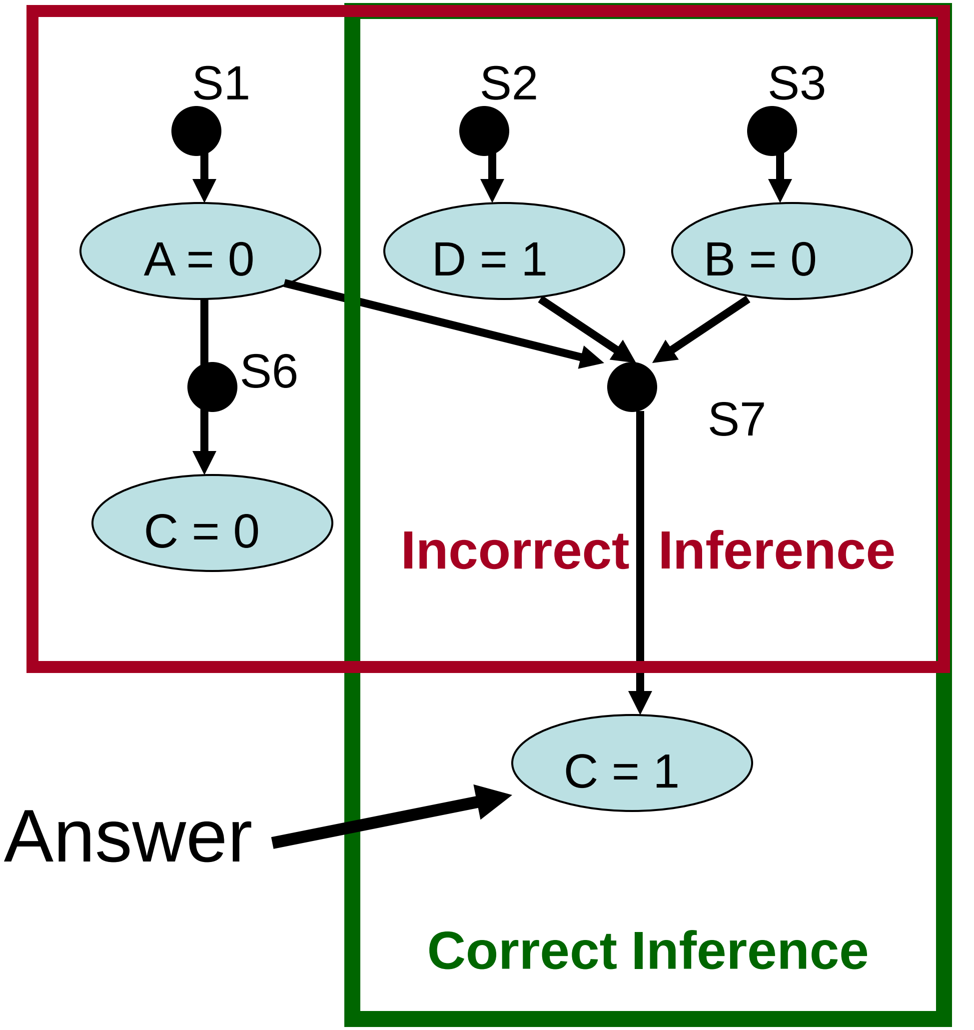
Correct Inference

- A **correct inference** for a test case is a complete state that contains the evidence, answer and has higher probability than any incorrect inference.
- An **incorrect inference** is a complete state that contains the evidence and a r.v. incompatible with the answer.



Force Correct Inference

- Looking at the example, if the **incorrect inference** has a <u>higher probability</u> then the **correct inference** one could try to adjust the weights of S-nodes **S6** and **S7** to reverse the probabilistic inequality.
- S-node weight changes are a non-structural modification, but are <u>not</u> always possible.

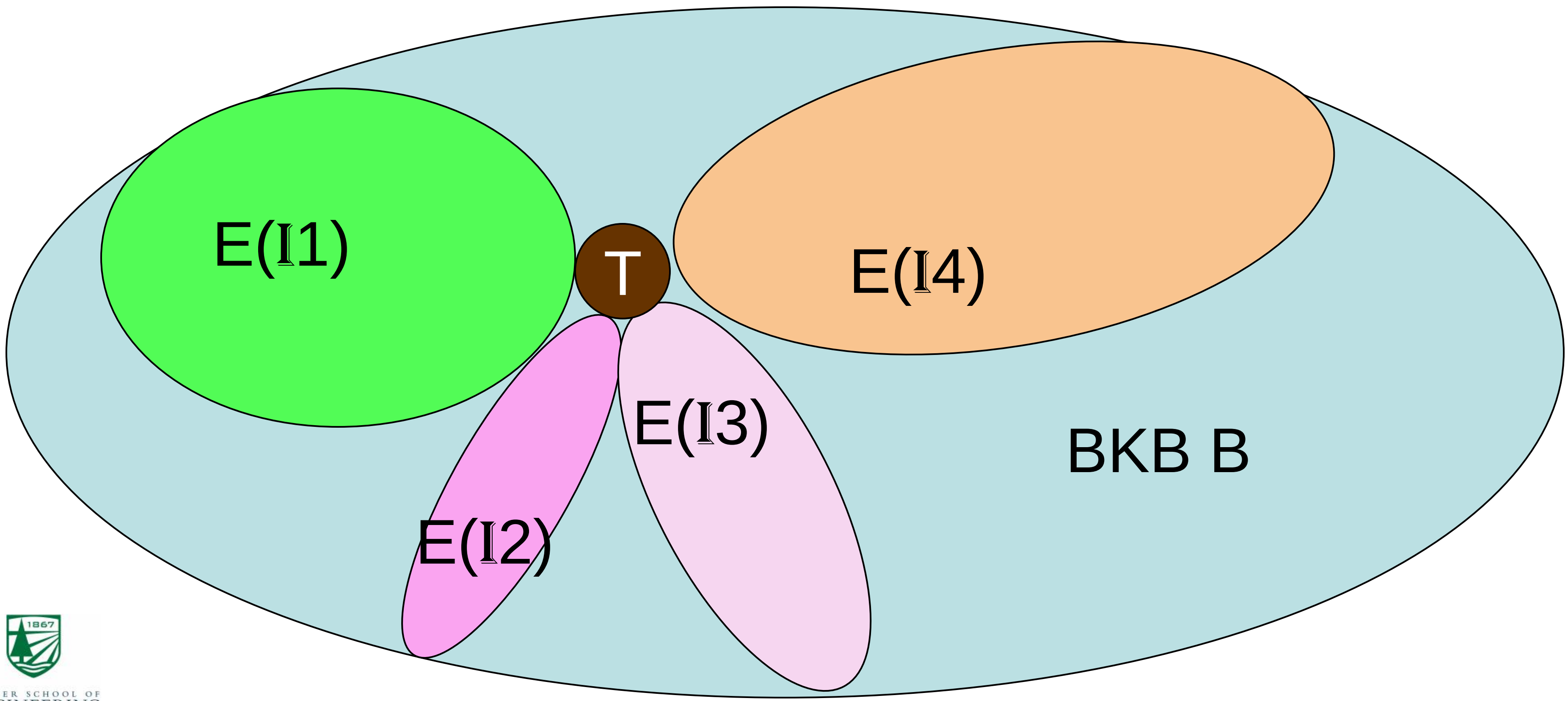


THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

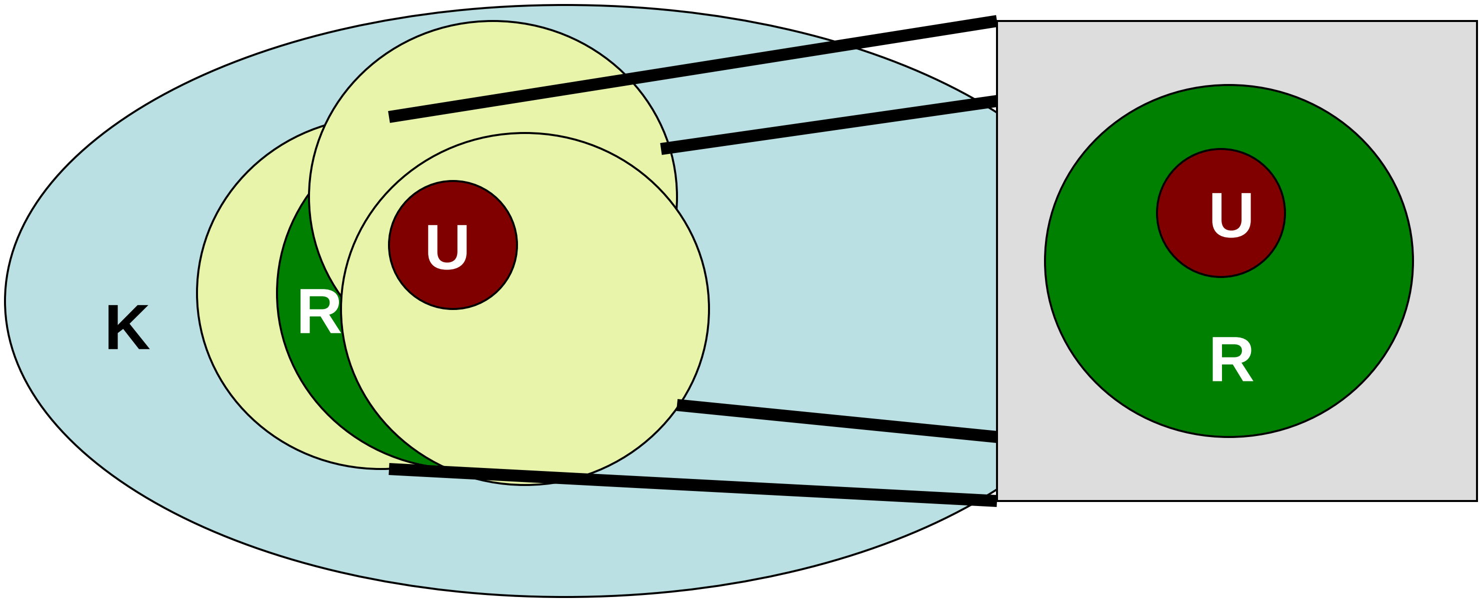
$\cup_{I \in I_B(T)} E(I) = Q(T)$ where $Q(T)$ is the set of all complete inferences in BKB B that are compatible with T

Also: $p(T) = \Sigma_{I \in Q(T)} P(I)$

Where $p(T)$ is the unique joint probability distribution $(B \models p)$ that is the sum of all complete inferences I over B



Reasoning w/ Inferences



Given evidence $\{u_1, \ldots, u_n\}$ and BKB K
Find inference $\{r_1, \ldots, r_m\}$ where n < m
$U \subset R$, $R \subseteq K$ and R is connected s.t.
$P(R|U)$ is maximized for all inferences
that meet the above condition.

Probability of an Inference

- The probability of an inference $\mathbb{I}$ is the probability that the state 'event' $\mathbb{I}$ happens which is the sum of probability of all complete states consistent with $\mathbb{I}$ {*the complete states to which* $\mathbb{I}$ *is relevant*}.

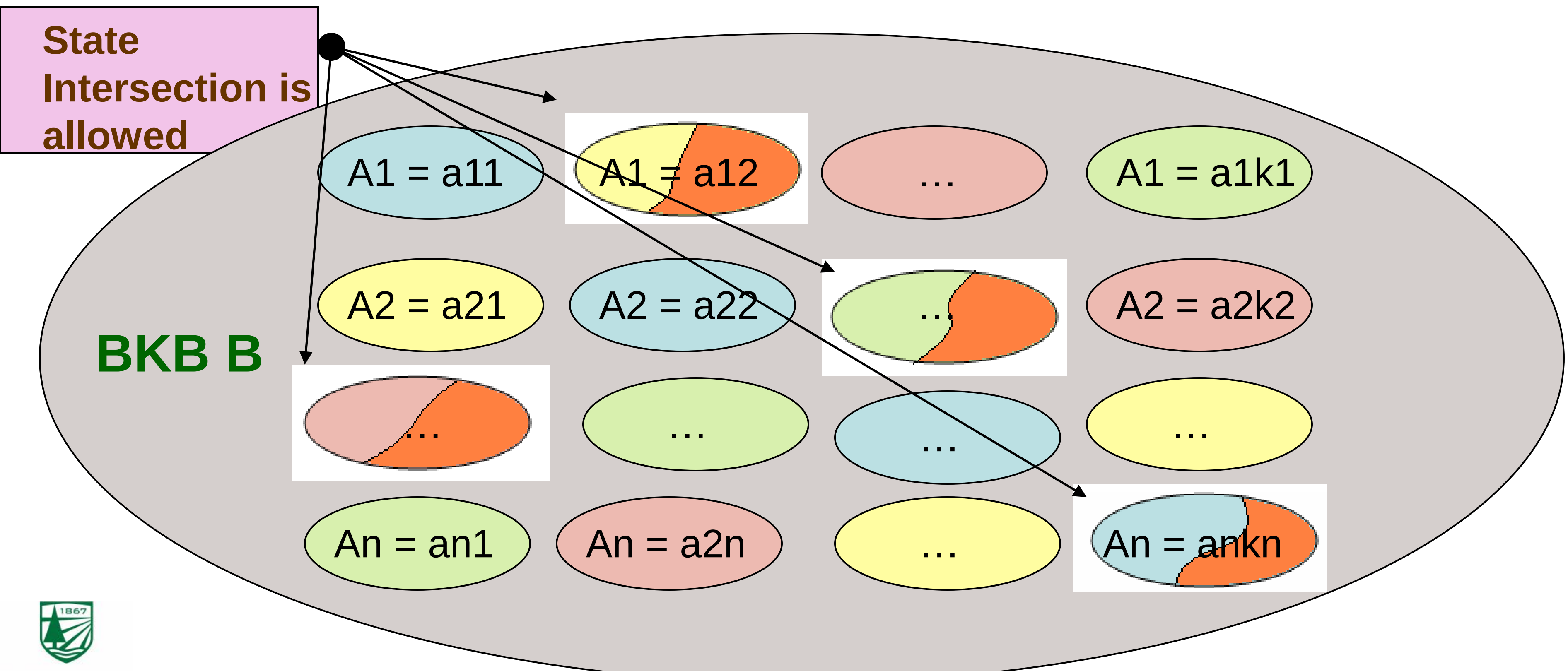
The joint probability of an inference τ of a BKB β is the product of its S-nodes

$$\Pi_{S \in S_\tau} P(S)$$

Assignment Complete: A BKB is assignment complete if for every complete state there exist a complete inference that represents that complete state.

Probability Complete: A BKB is probability complete if the sum of all complete inferences equals one.

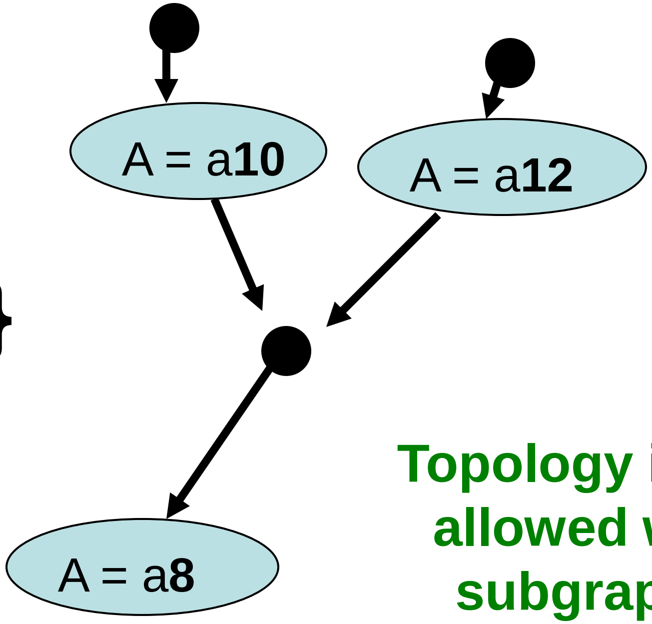
The **BKB** bellow would be **assignment complete** if the **Blue**, **Yellow** **Red**, **Green & Brown** complete states where subsumed by a complete inference AND would be **probability complete** if the probabilities of those inferences equaled one.



Topological Ordering

Let $\{c_1, \ldots, c_m\}$ be the I-nodes of an inference where $c_i \iff$ to $A_i = a_i$

Good: 8 < min{10,12}



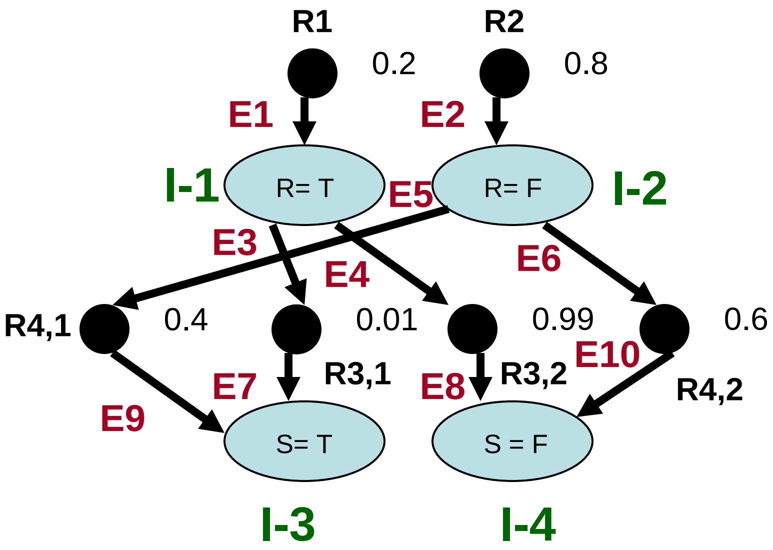
Topology is only allowed when subgraph is acyclic

*Then the **Probability** of the Inference would be:*

$$\prod_{i=1}^{m} P(A = a_i | A = a_{i+1}, \ldots, A = a_m)$$

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Topological Ordering for Quasi-Unique Representation



Depth First?

What do you do about Cyclicity?

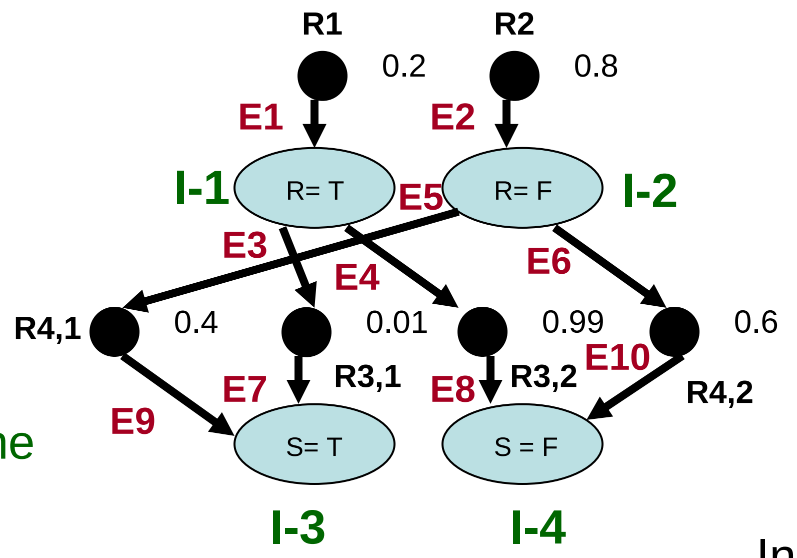
Have not formally decided …

On both vertices and arcs?

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Adjacency

1. Tailless S-nodes removed from columns.
2. Headless I-nodes removed from rows
3. S-node rows have exactly on element
4. I-node rows have one or multiple unity values



Incidence

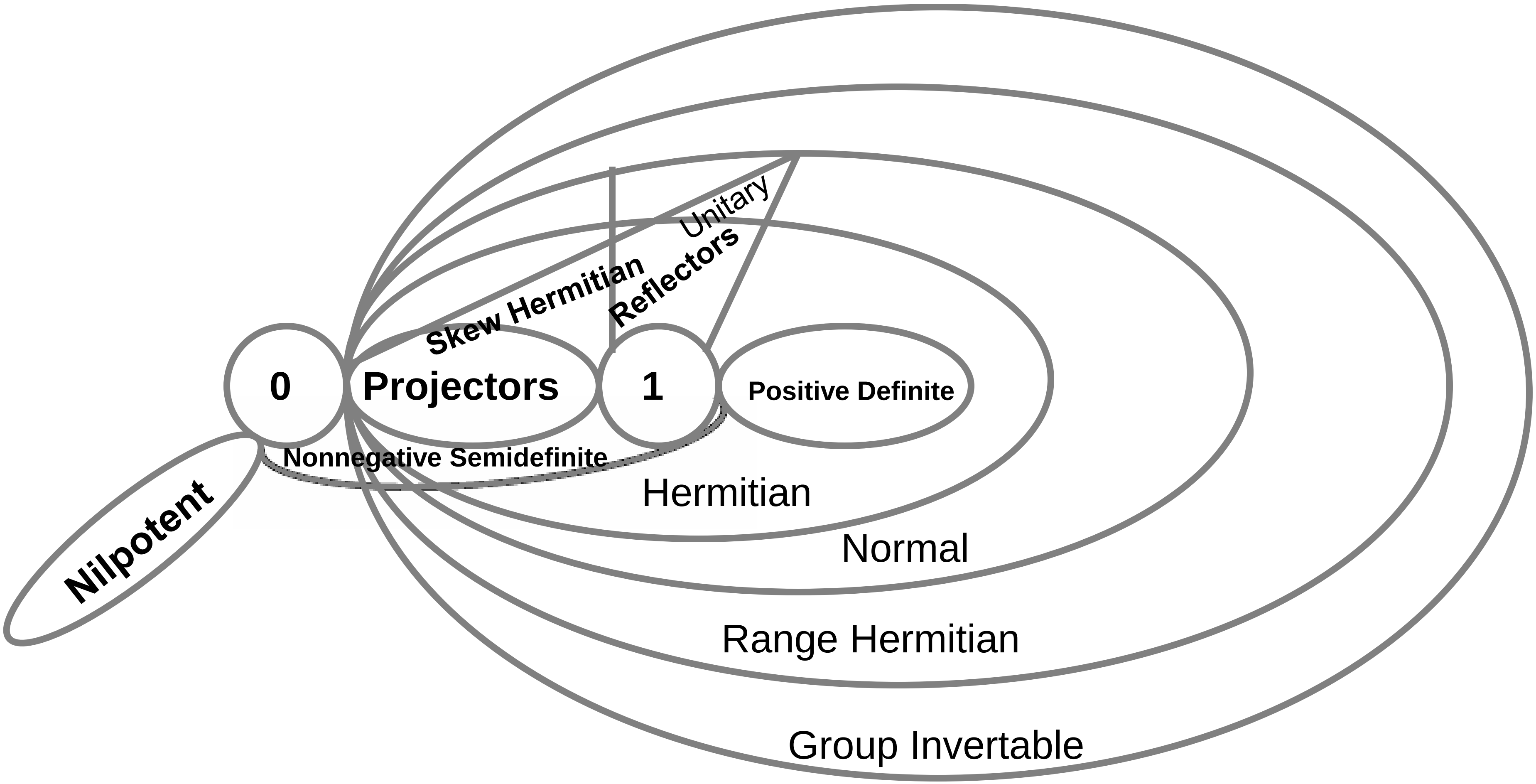
1. Two elements in every column: one positive one negative
2. For every row |negative| = # edges leaving, |positive| = # edges leaving

Adjacency Matrix

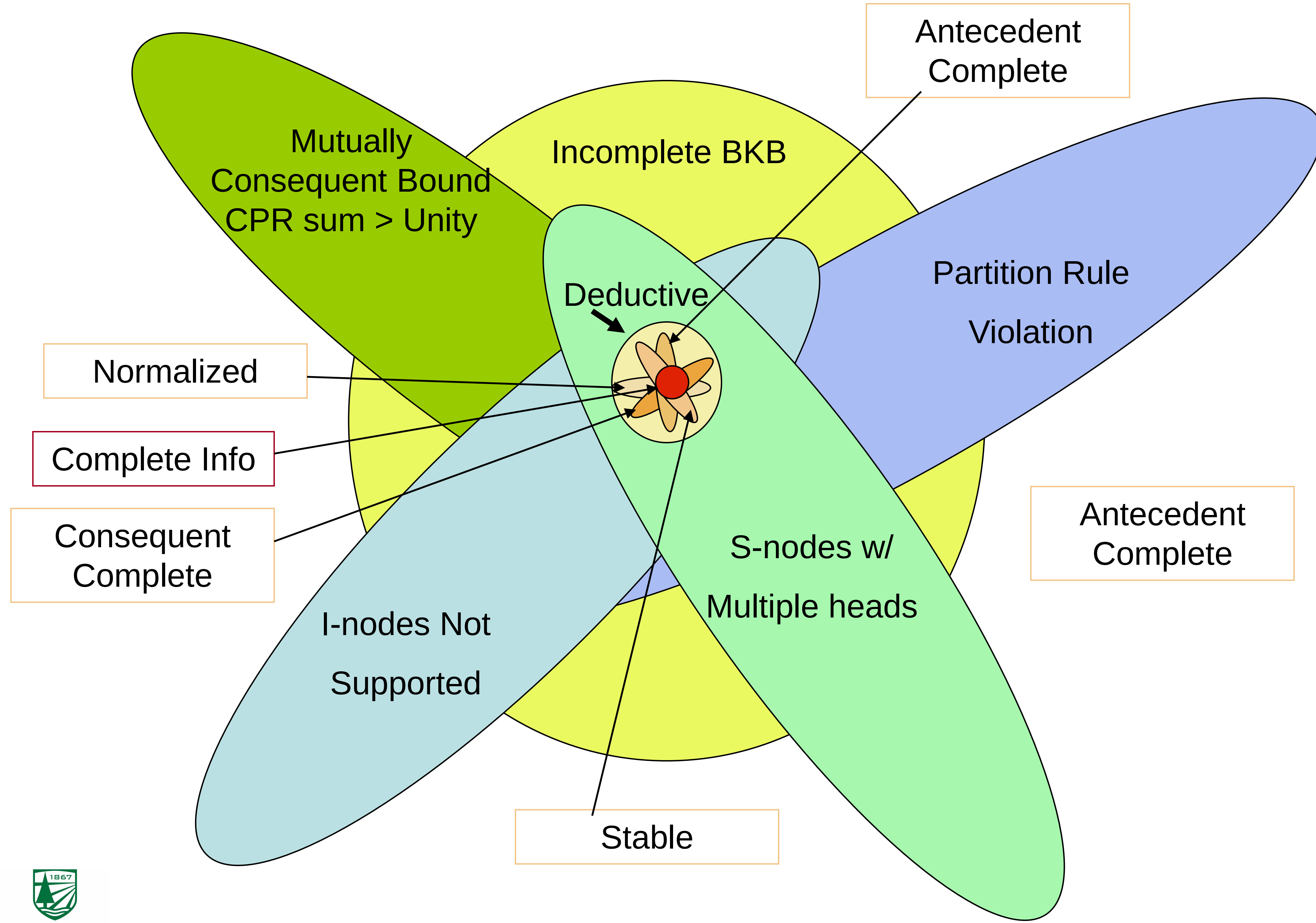
	I1	I2	R3,1	R3,2	R4,1	R4,2	I3	I4
R1	0.2							
I1			1	1				
R2		0.8						
I2					1	1		
R3,1							0.01	
R3,2								0.99
R4,1							0.4	
R4,2								0.6

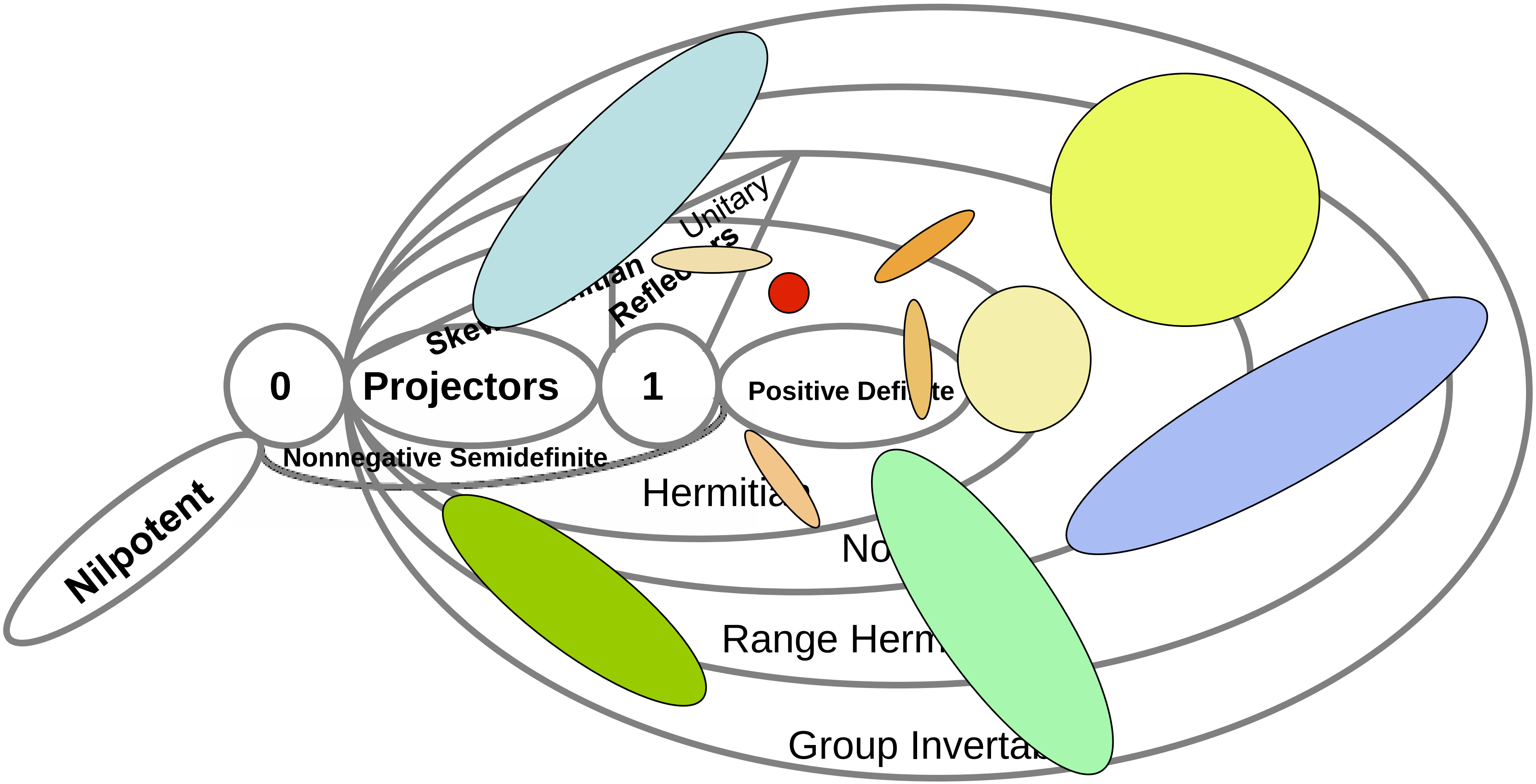
Incidence Matrix

	E1	**E2**	**E3**	**E4**	**E5**	**E6**	**E7**	**E8**	**E9**	**E10**
R1	-0.2	0	0	0	0	0	0	0	0	0
I1	0.2	0	-1	-1	0	0	0	0	0	0
R2	0	-0.8	0	0	0	0	0	0	0	0
I2	0	0.8	0	0	-1	-1	0	0	0	0
R3,1	0	0	1	0	0	0	-0.01	0	0	0
R3,2	0	0	0	1	0	0	0	-0.99	0	0
R4,1	0	0	0	0	1	0	0	0	-0.4	0
R4,2	0	0	0	0	0	1	0	0	0	-0.6
I3	0	0	0	0	0	0	0.01	0	0.4	0
14	0	0	0	0	0	0	0	0.99	0	0.6



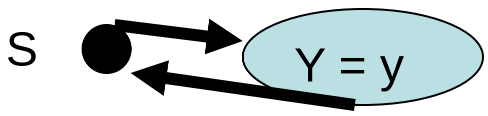
THAYER SCHOOL OF ENGINEERING AT DARTMOUTH


Antecedent Complete
Mutually Consequent Bound CPR sum > Unity
Incomplete BKB
Partition Rule Violation
Deductive
Normalized
Complete Info
Consequent Complete
Antecedent Complete
S-nodes w/ Multiple heads
I-nodes Not Supported
Stable



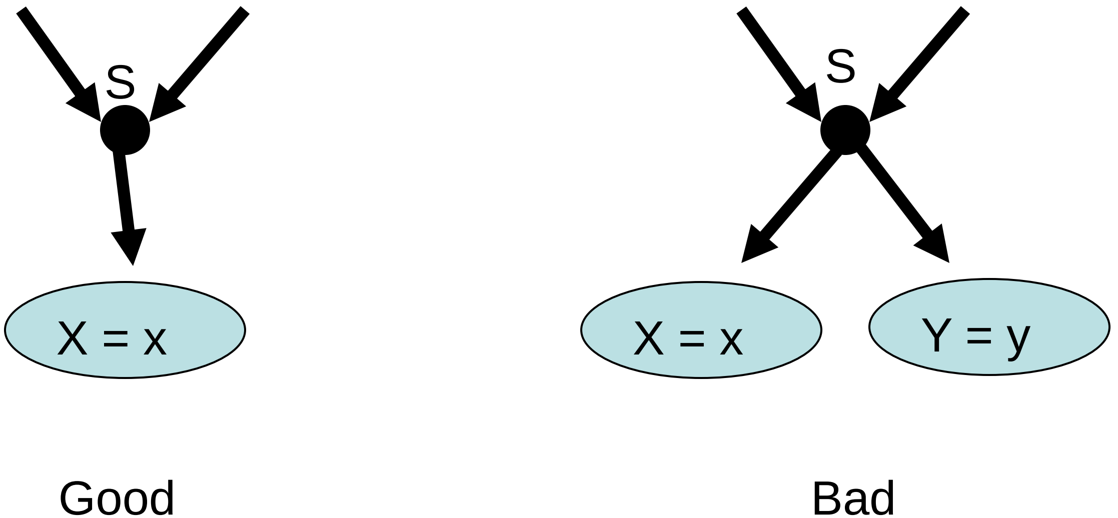
0
Projectors
1
Positive Defi te
Nonnegative Semidefinite
Unitary
Reflec rs
Skew itian
Hermitia
No
Range Herm
Group Inverta
Nilpotent

No self-pointing I-node



The antecedent and consequent of any CPR has null intersection

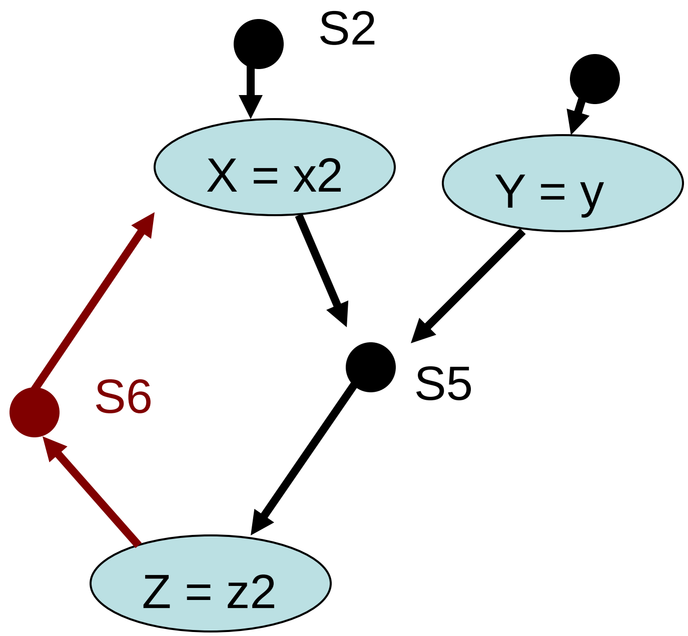
S-Nodes have exactly one I-Node Descendent



P(X = x, Y = y | Z = z , …) = S ➔ NOT ALLOWED

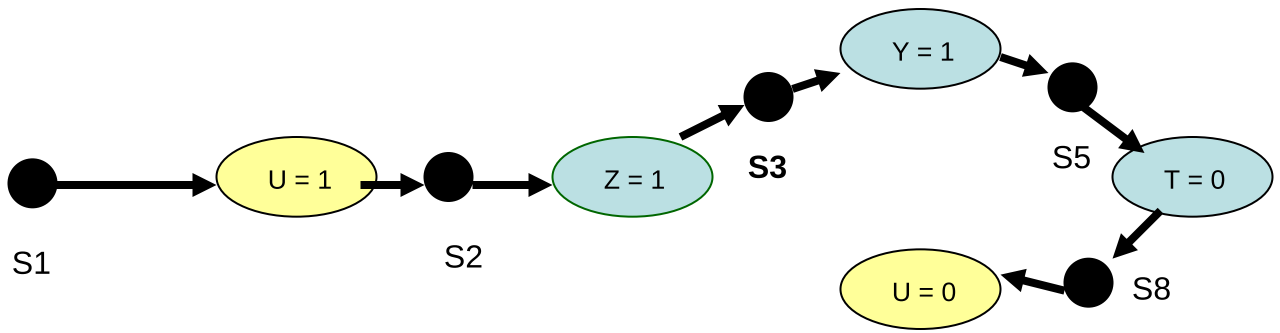
THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Inferencial Acyclicity



Acyclicity in an inference gets at the fact that no inference chain can have a support node head that occurs in the tail of any of its predecessors in that chain i.e. S6's head is at the tail of its predecessor S5

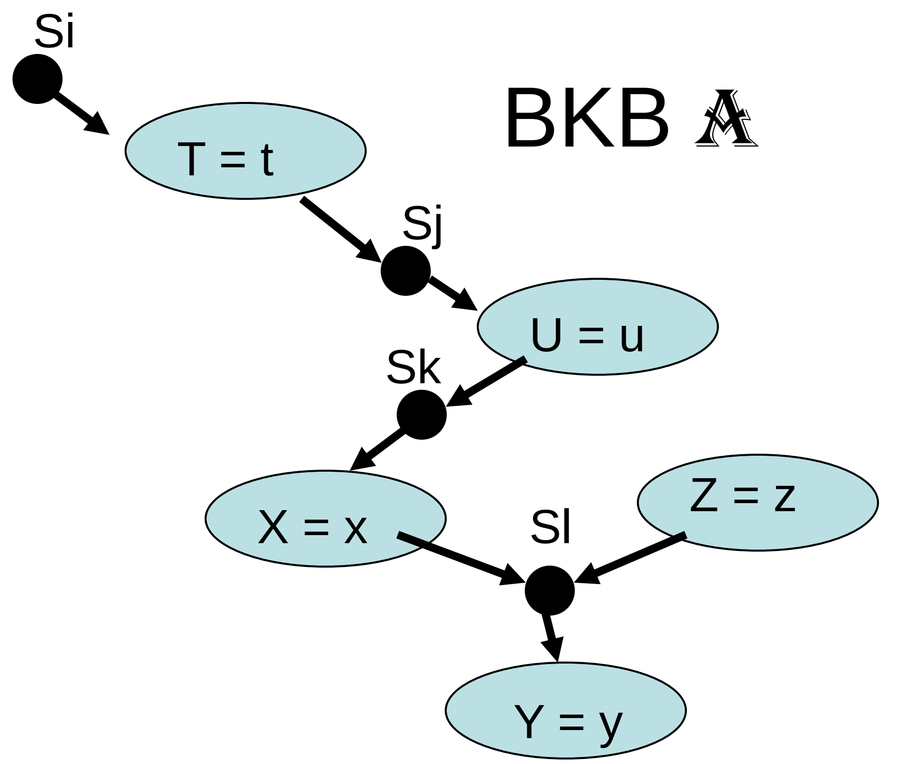
Variable Cycle



A variable cycle in BKB is a directed path that includes two I-nodes (*not necessarily distinct*) from the same r.v. A cycle that starts and ends at the same I-node counts that I-node twice and is a variable cycle.

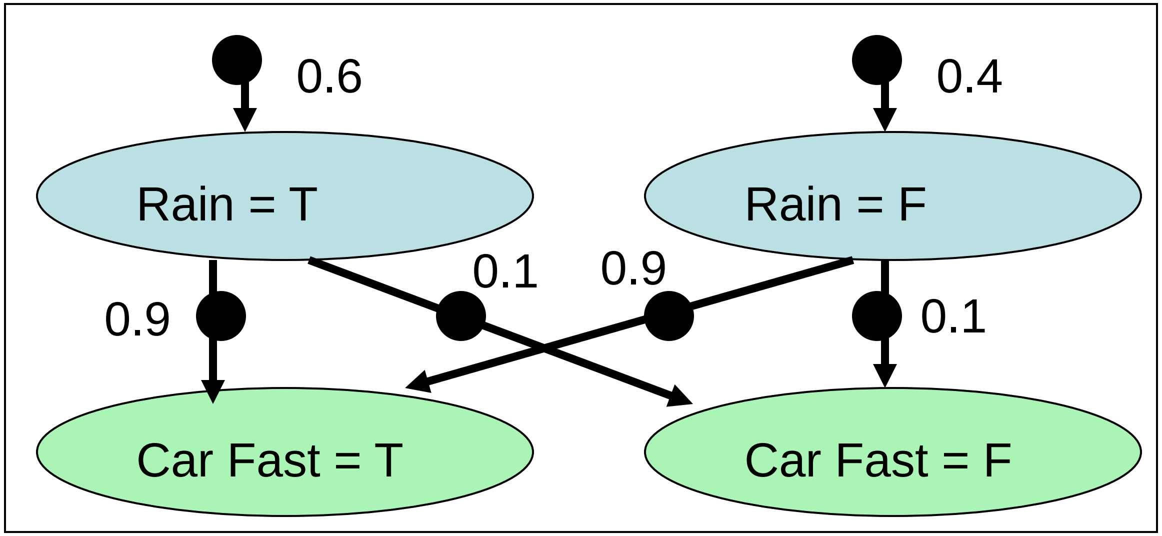
Every BKB has to respect a weighting

- All S-nodes are conditional probability rules that take values in [0,1]
- All BKBs $\mathfrak{A}$ have to obey a weighting $\mathbb{W}$ that is the product of its S-nodes



Here $\mathbb{W}(\mathfrak{A})$ = P(T = t)*P(U = u | T = t)*P(X = x | U = u)*P(Y = y | X = x, Z = z)

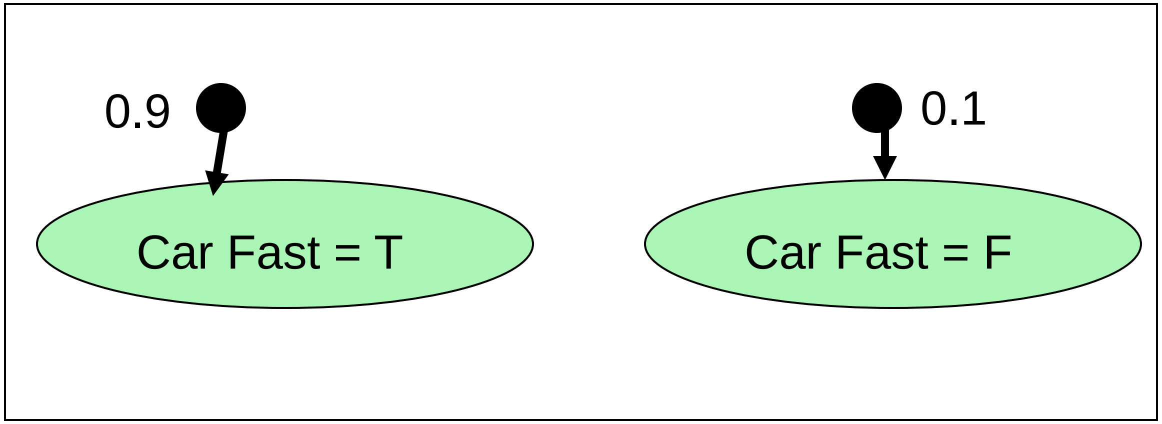
Node Removal?



Graph A

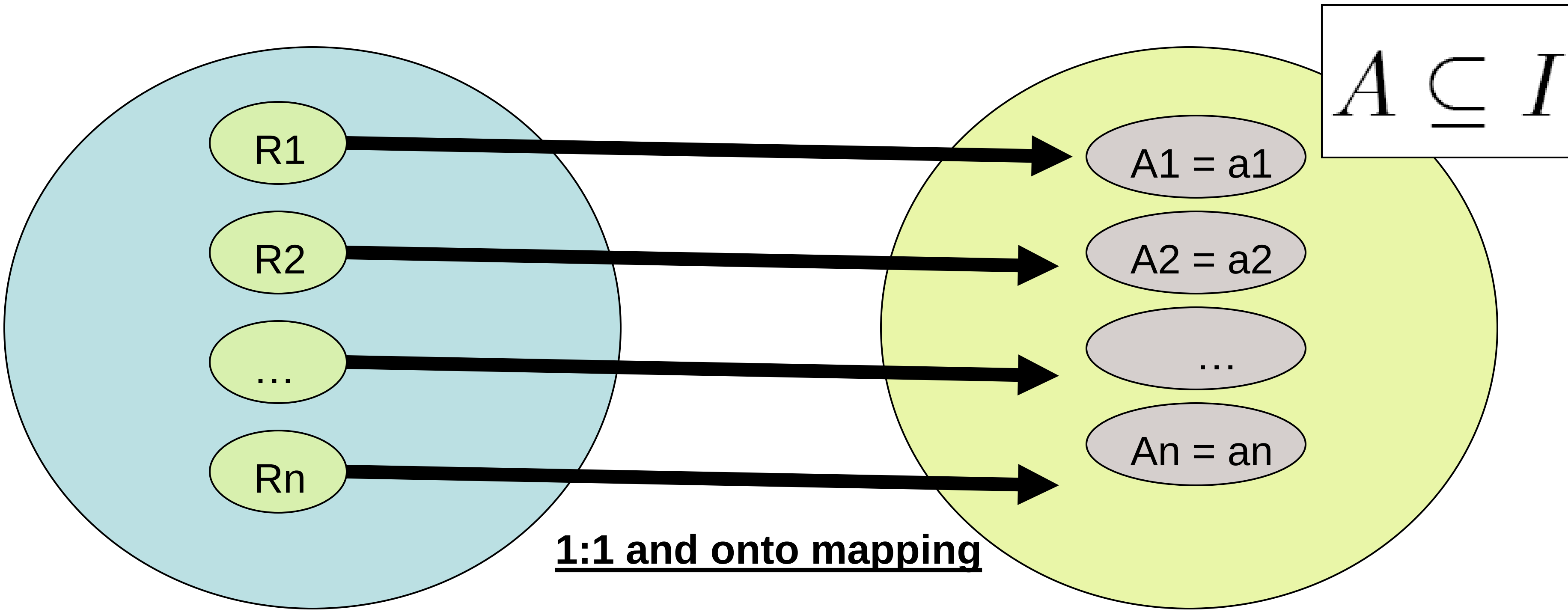
You can always remove r.v.s. but the distributions will **not** be the same. In the example bellow, the S-nodes of Graph B are the derived marginal distributions from Graph A's joint distribution.

Graph B



THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Deductive Subsets of CPRs



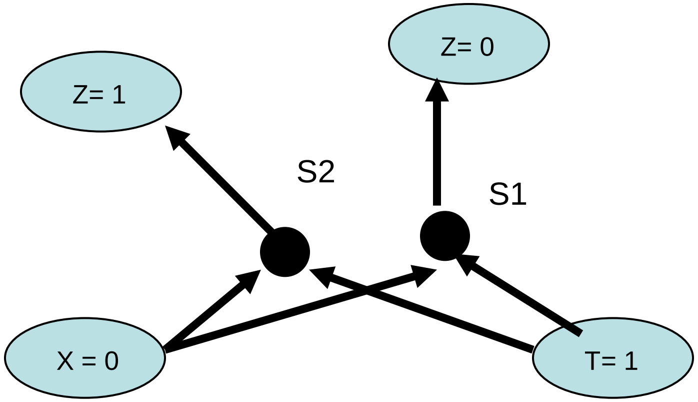
$S \subseteq K$

A subset S of a BKB K $S \subseteq K$ is called deductive
if for each CPR $R \in S$ where:
$R : A_{i1} = a_{i1} \wedge A_{i2} = a_{i2} \wedge \ldots \wedge A_{in-1} = a_{in-1} \Longrightarrow A_{in} = a_{in}$
and for each $k = 1, 2, \ldots, n : \exists R_k \in S$ s.t.
$con(R_k) = \{A_{ik} = a_{ik}\}$
Also $\nexists R' \epsilon S$ where $R' \neq R$ and $con(R') = con(R)$
in otherwords $\Sigma_{R \in S} con(R)$ spans $\{A_1, A_2, \ldots, A_n\}$
and $A_i = a \wedge b \Leftrightarrow a = b$

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Consequent Variant

A graph is called consequent-complete if every maximal C-variant set is also complete



Note: Consequent Variant is different then Consequent Bound

A set of CPRs are consequent variant if they have the same set of antecedents and consequent random variable. Here {S1,S2} are consequent variant because they share the same consequent r.v. Z, and both have the same set of antecedent I-nodes namely {X = 0 and T = 1}, additionally if 0 & 1 are the only instantiations of r.v. Z then {S1,S2} would be consequent variant complete for r.v. Z.

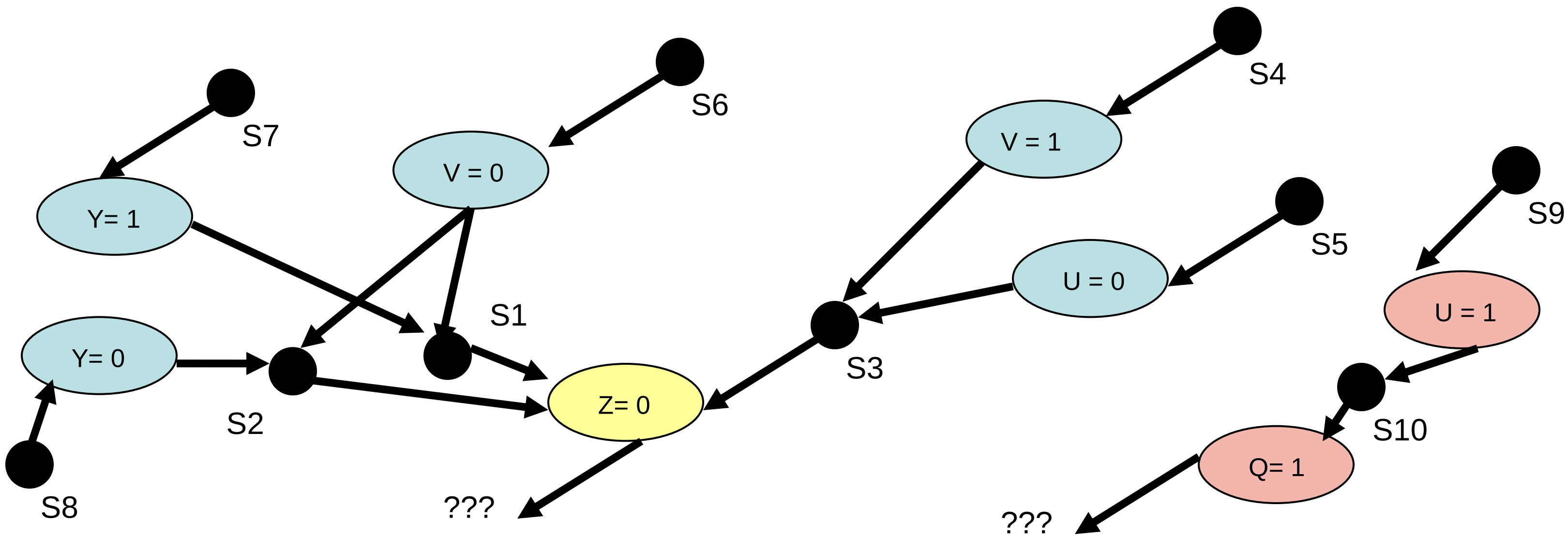
THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Consequent Completeness

- Consequent Completeness means that if there is a rule that can deduce an I-node {instantiation of r.v. X} from some antecedent state, then all other instantiations of X may be deduced from the same antecedent state.
- A BKB is **consequent-complete** if every maximal C-variant set is also complete

Antecedent Variant

The pink annex is INCONSISTENT with antecedent-cover

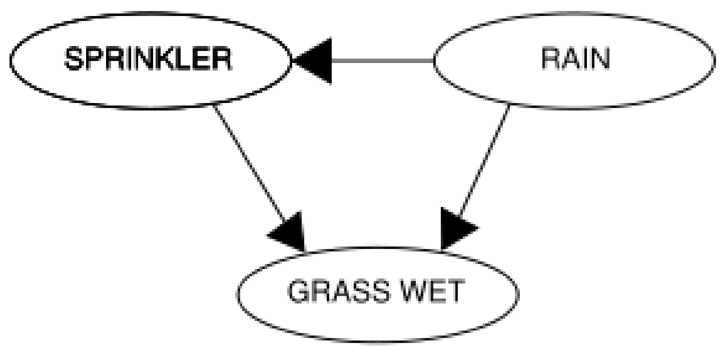


A set of CPRs ℟ is called **antecedent-variant** if they have the same consequent I-node. Additionally, ℟ is called a **cover** of its antecedent-variables VA if all possible states for VA are **CONSISTENT** with the antecedent of some rule in ℟. If all possible states for VA are also equal to the antecedent of some rule in ℟, then ℟ is **antecedent-compete** for VA. i.e. the above BKB K has CPRs ℟ = {S1,S2,S3}, VA = **{Y,V,U},** ante(℟) = **{Y=0,Y=1,V=0,V=1,U=0},** if no state on VA exist outside of ante(℟) then ℟ would be an **antecedent-cover**. If ante(℟) = **{Y=0,Y=1,V=0,V=1,U=0,U=1}** it would be a complete enumeration on VA and thus ℟ would be **antecedent-complete**. A graph is **antecedent-cover/compete** if every maximal **antecedent-variant** are a **cover/complete**.

BN correspond to BKB?

- *A Bayes Network corresponds naturally to a BKB that is acyclic, consequent-complete and antecedent-complete.*

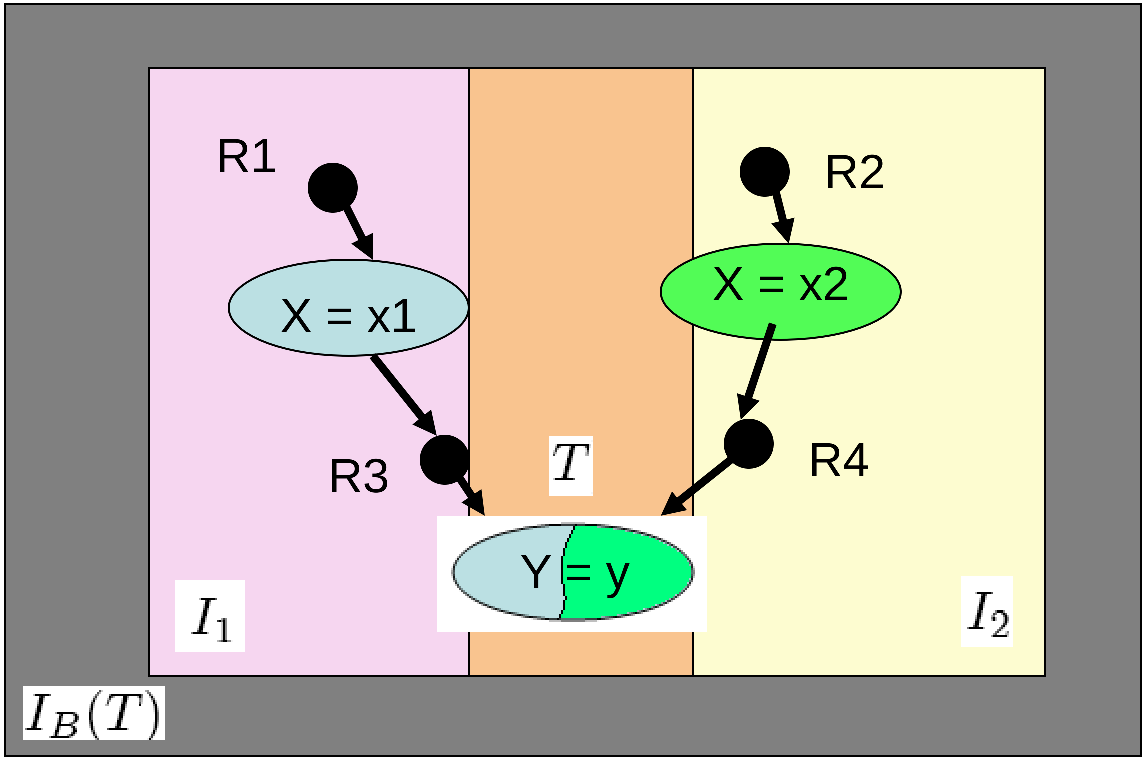
	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



RAIN	
T	F
0.2	0.8

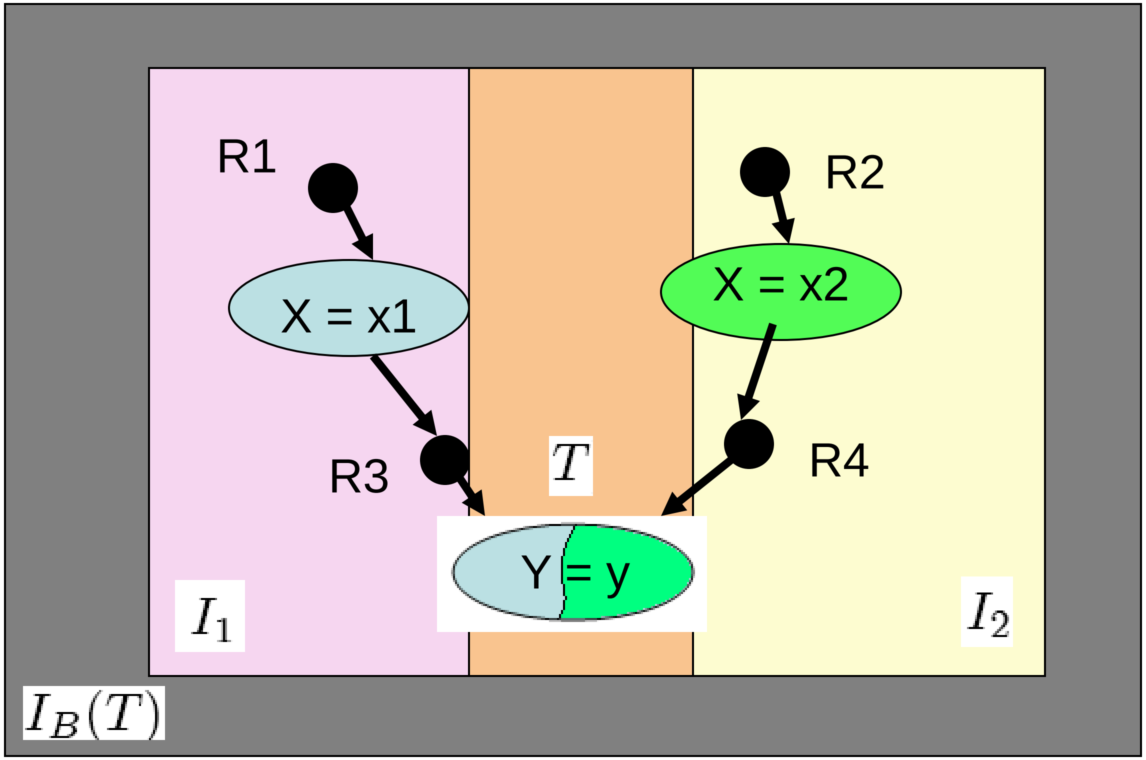
		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

A BKB with these conditions has its consistency trivially meet.



$I_B(T)$ is the set of all inferences where:
$\forall I \in I_B(T)$ space $\Rightarrow \Sigma_{R \in I} con(R) \geqq V(T) \wedge \nexists\, I' \subset I$ where I' is minimal w.r.t. T
Where $V(T)$ is the set of all I-nodes in T
Additionaly, two discrete inferences I_1 and I_2 in $I_B(T)$
will have to be mutex i.e. "incompatible"

THAYER SCHOOL OF
ENGINEERING
AT DARTMOUTH

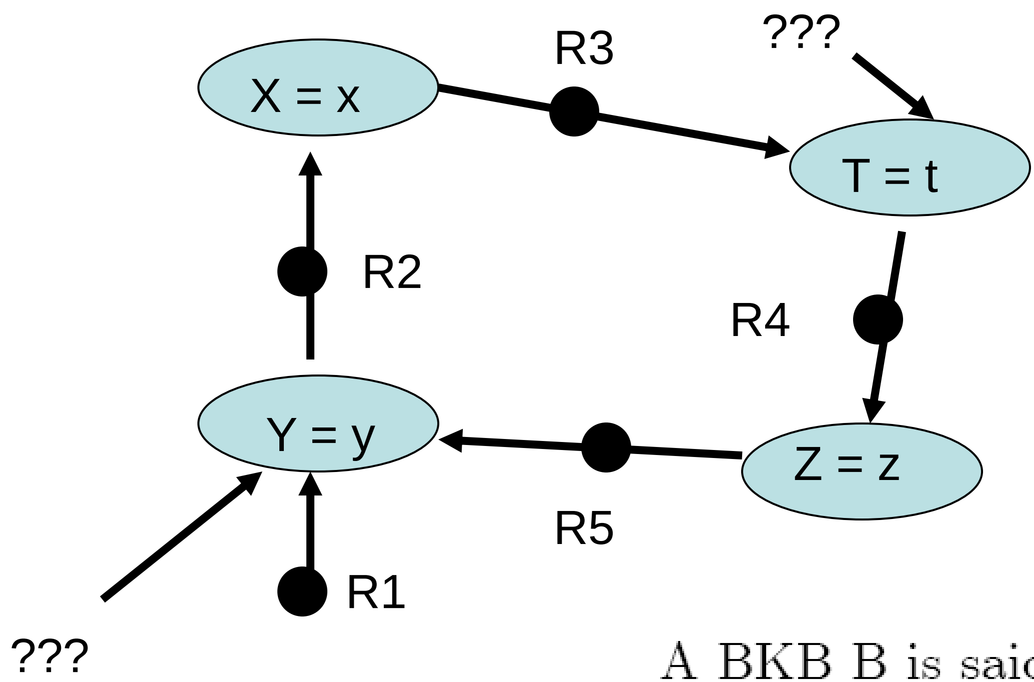


Let E(I) denote the set of all complete inferences that are supersets of I. For $I_1, I_2 \in I_B(T)$ where $I_1 \neq I_2$
$E(I_1) \cap E(I_2) = \phi$ in the example above
$E(I_1) = I_1$ and $E(I_2) = I_2 \Rightarrow$ their intersection is an unsupported I-node which is not an inference.

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Stable BKB B

Subset S



A BKB B is said to be stable if all subsets $S \subseteq B$ of the form $S = \{R_1, R_2, ..., R_n\}$ where if $con(R_n) = \{A = a\} \nexists \{A = a'\} \in ant(R_1)$

a need not be distinct from a'

The example to the left demonstrates that stability need not exclude directed cycles

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

BKB B is conditional probability consistent with p when

A BKB $B \models p$ if $P(R) \models p \; \forall R \in B$
and that happens when B is stable, assignment complete and probability complete

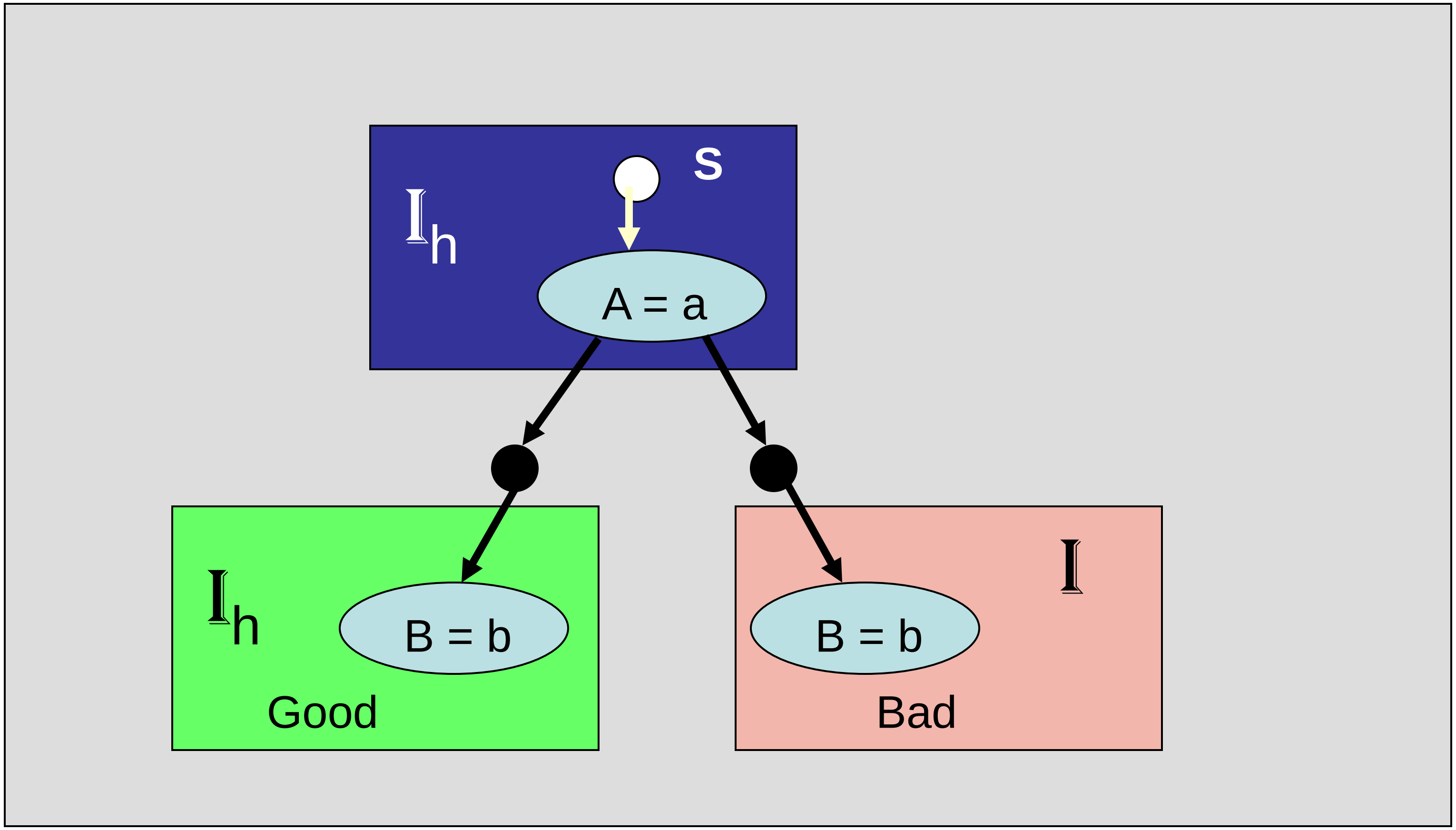
Given an inference Γ & subgraph H of G

$\Gamma = \{S, I, E\}, H = \{S_h, I_h, E_h\}$

Where

$H \subseteq G \vee \Gamma \subseteq H$

If $a \epsilon I_h \nexists b \epsilon DEC(a)$ where $b \epsilon I_h \backslash I$



Importance Sampling

- Complicated Belief heuristic for determining the most probable **complete state** $\mathbb{S}$ and associated sample weight $\mathbb{W}(\mathbb{S})$ for a grounded and normalized BKB $\mathbb{K}$ w/ cycles; that is randomized by a random uniform probability for each of the r.v.s. not in the evidence and a weight accumulator.

BKB $\mathbf{K} \rightarrow$ $R^1_1, \ldots, R^1_{k1}$, … , $R^n_1, \ldots, R^n_{kn}$

On set of r.v.s. $\mathbf{X} \rightarrow \{X1, \ldots, Xn\}$

Where $Cons(R^i_1) + \ldots, + Cons(R^i_{ki}) \rightarrow$ Domain Xi

Evidence $\mathbf{E} \rightarrow$ **X1, … , Xm** $0 <= m < n$

Initial Inference & associated complementary CPRs all initialized to null

$\mathbf{I} = \text{null}$

$R1$ $R2$ . . . Rn

While randomly generated uniform probability condition is not violated

$\mathbf{R^i_1, \ldots, R^i_{ki}}$

For all CPR set $\mathbf{R^i_1, \ldots, R^i_{ki}}$ for **i <= n** if $\mathbf{R^i_j}$ extends $\mathbb{I}$ w.r.t. **X**i then $\mathbf{R^i_j}$ is added to $\mathbb{R}$i

For all CPR set $\mathbf{R^i_1, \ldots, R^i_{ki}}$ for **i > m** if there exits CPRs **R** in $\mathbb{R}$i that extends $\mathbb{I}$ w.r.t. **X**i then $\mathbb{I}$ is unioned with **R**

After $\mathbb{I}$ is expanded $\mathbb{S}$ is determined to be st($\mathbb{I}$) unioned w/ $\mathbb{E}$ where st($\mathbb{I}$) is the smallest state that $\mathbb{I}$ is the MRI to

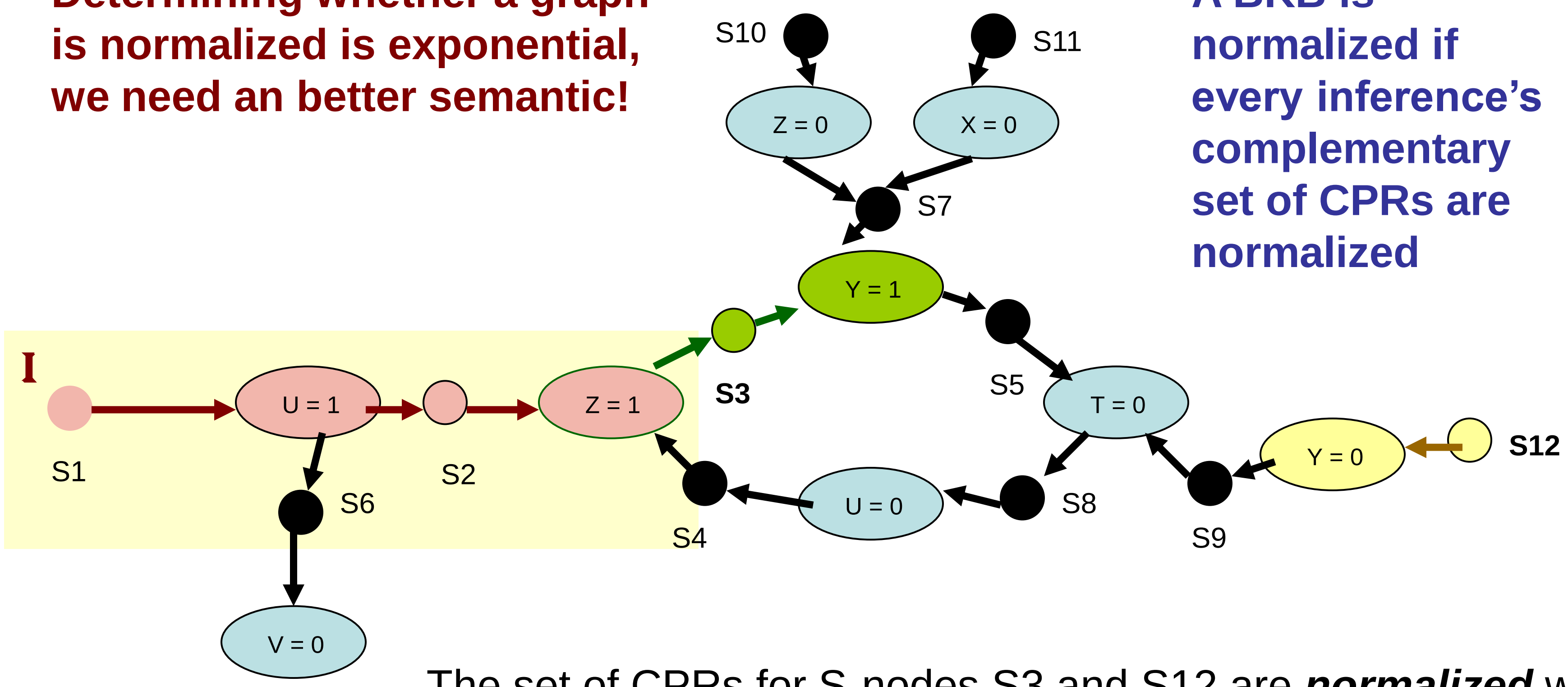
What Kind of BKB Makes Sense?

- If the BKB is **normalized**, **consequent-complete**, all nodes are **grounded**, all maximal inferences are **complete**. Then the BKB is **locally complete** w.r.t every rule $\mathbb{R}$ and $\mathbb{P}(\text{cons}(\mathbb{R})|\text{ant}(\mathbb{R})) = \mathbb{W}(\mathbb{R})$ whenever $\mathbb{P}(\text{ant}(\mathbb{R})) > \mathbb{0}$ ☺
- *Note: given a state $\mathbb{S}$ the **maximal relevant inference** is the largest inference relevant to $\mathbb{S}$ and the only one with weight greater than zero.*

A complementary set of CRPs is ***normalized*** w.r.t an inference and r.v. if the sum of its respective S-nodes are less than or equal to unity

Determining whether a graph is normalized is exponential, we need an better semantic!

A BKB is normalized if every inference's complementary set of CPRs are normalized



The set of CPRs for S-nodes S3 and S12 are ***normalized*** w.r.t the inference **I** and random variable Y iff S3 + S12 <= unity (*compete if = unity*). A graph is called normalized if for each inference its respective complementary set of CPRs is normalized.

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

CPR Weight and Significance?

- The weight of a conditional probability rule only has semantics in the context of its inference.
- Weights for ungrounded rules are meaningless

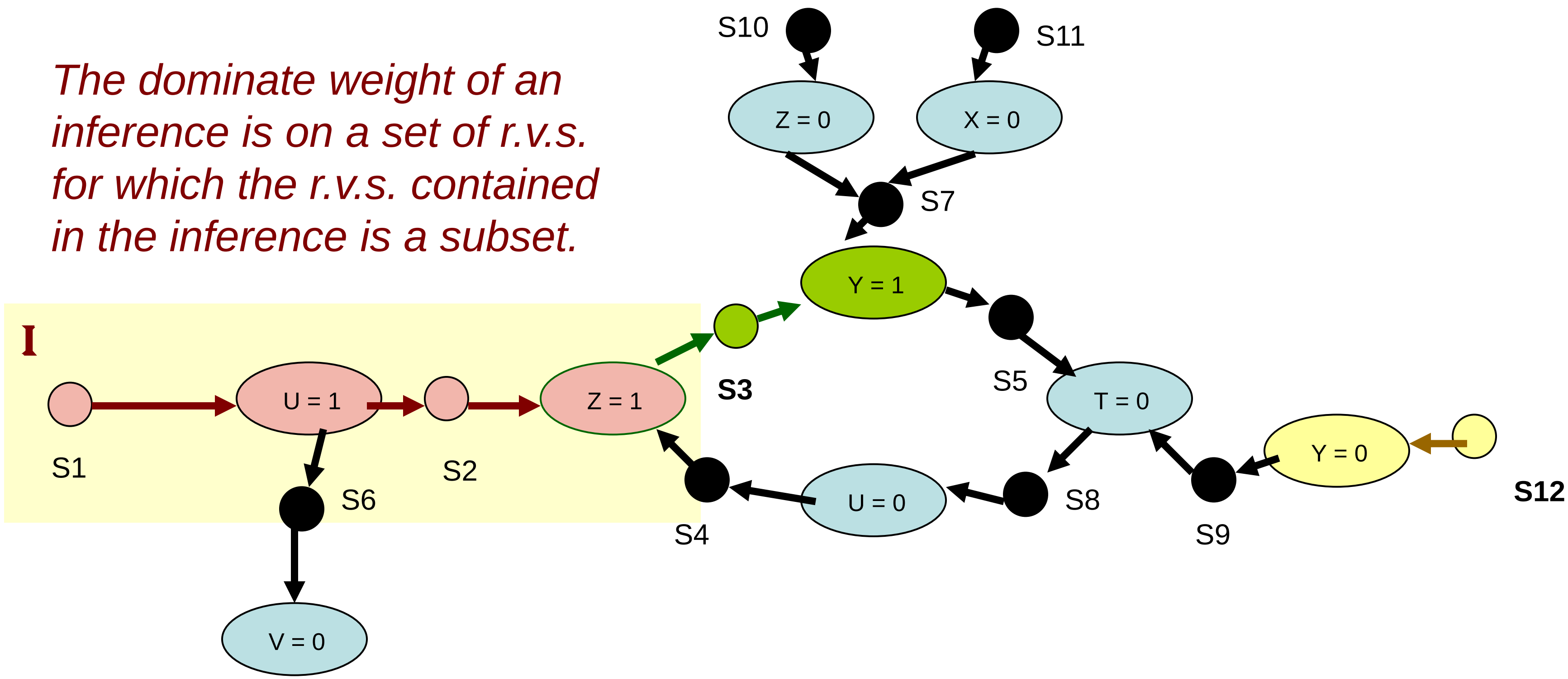
BKB Normalization

- Normalization helps simplify heuristics
- If $\mathbb{K}$ a consequent-complete BKB and each C-variant set of CPRs $\mathbb{R}$ is locally normalized {*$\mathbb{R}$ is complementary w.r.t. some inference and r.v. $\mathbb{X}$ where $\mathbb{X}$ is the consequent for $\mathbb{R}$*} then $\mathbb{K}$ is normalized.
- If all the nodes are grounded the above **sufficient** condition becomes **necessary**

1867
THAYER SCHOOL OF
ENGINEERING
AT DARTMOUTH

The ***dominate weight*** of an inference is the weight of the inference multiplied by the product of the positive difference between unity and the weight of the maximal complete set of CPRs for each r.v. not represented in the inference

The dominate weight of an inference is on a set of r.v.s. for which the r.v.s. contained in the inference is a subset.



For example the dominate weight for the inference I is:

S1S2[(1-S3-S12)(1-S11)(1-S6)]

When is a <u>function</u> $\mathfrak{F}$ consistent with a normalized BKB over a set of r.v.s.?

- Also note that if $\mathfrak{F}$ is consistent with a normalized BKB, then for any inference in the BKB the product of the S-nodes is identical to $\mathfrak{F}$ acting on the state of the inference.

When is a <u>function</u> $\mathfrak{F}$ consistent with a normalized BKB over a set of r.v.s.?

- When given any inference in the BKB the sum of $\mathfrak{F}$ across all complete states, ***in the set of complete states for which the inference is the maximal relevant inference***, is equivalent to the **dominate weight** of an inference. And is called the ***default distribution*** for the BKB if $\mathfrak{F}$ returns the same result across all the complete states it acts upon.

If $\mathbb{K}$ a normalized BKB over r.v.s $\mathbb{X}$, and $\mathbb{F}$ a distribution consistent w/ $\mathbb{K}$ then $\mathbb{F}$ is a joint probability distribution over the set of complete states on $\mathbb{X}$

- If the set of r.v.s. is $\{\mathbb{U},\mathbb{V},\mathbb{Z}\}$ each with domain {0,1} then the cross product of the domains is the set of ordered triplets of 0s and 1s which is the same thing as the set of complete states on $\mathbb{X}$, therefore the domain of $\mathbb{F}$ is the cross-product of all the domains in $\mathbb{X}$, of whose complete enumeration where evaluate would evaluate to unity.

THAYER SCHOOL OF ENGINEERING AT DARTMOUTH

Is this an Inference?

Rule Violation: there are two S-nodes pointing to one I-node that do not share a r.v. w/ distinct instantiation; therefore this graph is not mutually exclusive w.r.t. a partition rule

But: There is a chance, for this subgraph, that all the S-nodes are well-defined and well-founded, the I-nodes are well-supported and it is acyclic; ***what is the difficulty here outside of the partition rule?***