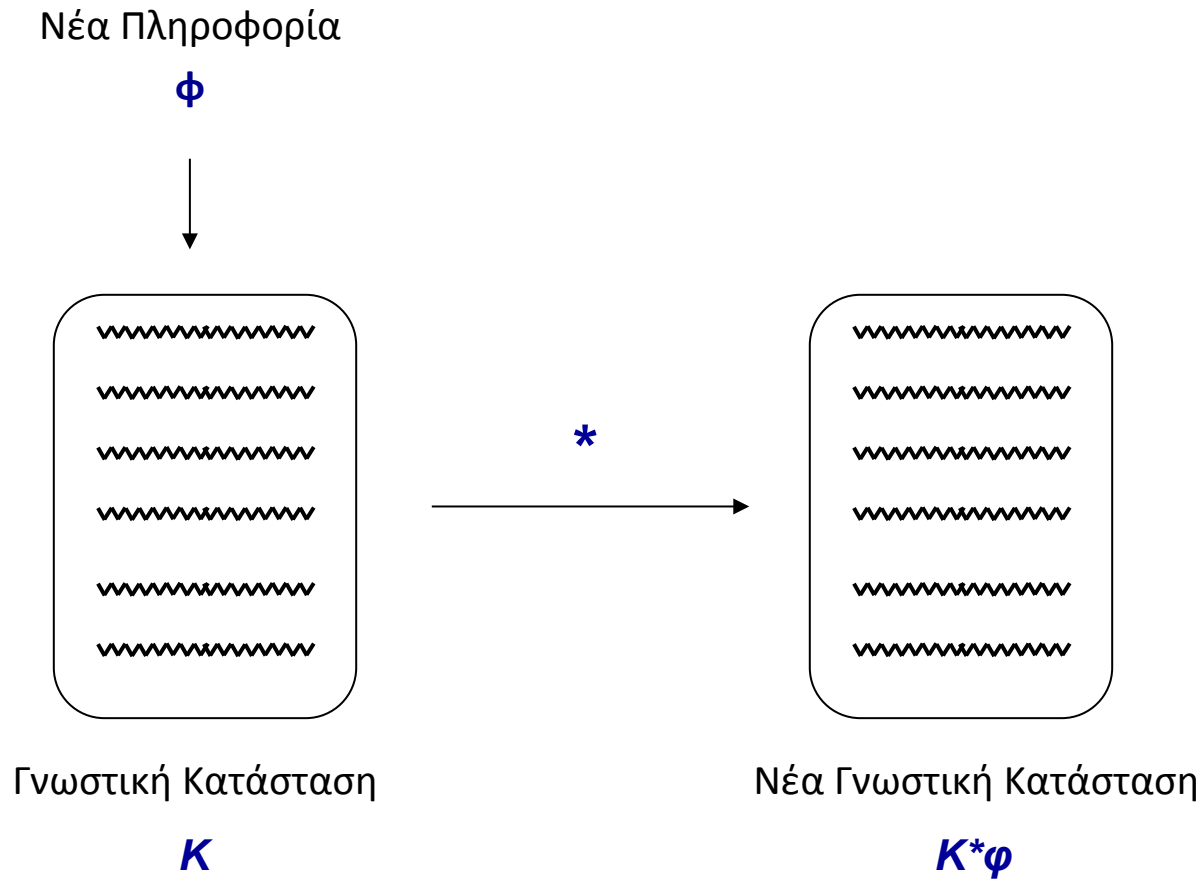


Τεχνητή Νοημοσύνη II

Παύλος Πέππας

Τμήμα Ηλεκτρολόγων Μηχανικών
και Τεχνολογίας Υπολογιστών

Αναθεώρηση Πεποιθήσεων



Παράδειγμα

Το αμάξι του Γιάννη είναι BMM.

Τα BMM κατασκευάζονται στην Γαλλία

Η Γαλλία ανήκει στην ΕΕ.

Όλα αμάξια κατασκευάζονται στην ΕΕ παίρνουν αμόλυβδη.

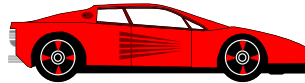
Το αμάξι του Γιάννη παίρνει αμόλυβδη.



Παράδειγμα

Το αμάξι του Γιάννη είναι BMM.
Τα BMM κατασκευάζονται στην Γαλλία
Η Γαλλία ανήκει στην ΕΕ.
Όλα αμάξια κατασκευάζονται στην ΕΕ παίρνουν αμόλυβδη.

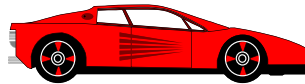
Το αμάξι του Γιάννη παίρνει αμόλυβδη.



Παράδειγμα

Το αμάξι του Γιάννη είναι BMM.
Τα BMM κατασκευάζονται στην Γαλλία
Η Γαλλία ανήκει στην ΕΕ.
Όλα αμάξια κατασκευάζονται στην ΕΕ παίρνουν αμόλυβδη.

Το αμάξι του Γιάννη παίρνει αμόλυβδη.



Ορθολογική
Αναθεώρηση
Πεποιθήσεων

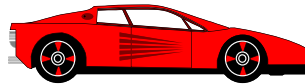
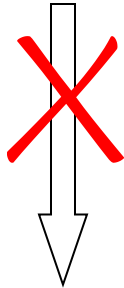
Τα BMM κατασκευάζονται στην Γαλλία
Η Γαλλία ανήκει στην ΕΕ.
Όλα αμάξια κατασκευάζονται στην ΕΕ παίρνουν αμόλυβδη.
Το αμάξι του Γιάννη δεν παίρνει αμόλυβδη.



Παράδειγμα

Το αμάξι του Γιάννη είναι BMM.
Τα BMM κατασκευάζονται στην Γαλλία
Η Γαλλία ανήκει στην ΕΕ.
Όλα αμάξια κατασκευάζονται στην ΕΕ παίρνουν αμόλυβδη.

Το αμάξι του Γιάννη παίρνει αμόλυβδη.



**Το αμάξι του Γιάννη δεν παίρνει
αμόλυβδη.**
Υπάρχει ζωή στον Άρη.



Τα BMM κατασκευάζονται στην Γαλλία
Η Γαλλία ανήκει στην ΕΕ.
Όλα αμάξια κατασκευάζονται στην ΕΕ παίρνουν αμόλυβδη.
Το αμάξι του Γιάννη δεν παίρνει αμόλυβδη.



Ορθολογική
Αναθεώρηση
Πεποιθήσεων

Πως Ξεκίνησε

(Carlos Alchourron and David Makinson)

Παλιός Νομικός Κώδικας

- Τα παιδιά μπορούν να δουν TV μόνο αν έχουν φάει βραδινό.
- Τα παιδιά μπορούν να φάνε βραδινό μόνο αφού τελειώσουν το διάβασμα.

Πως Ξεκίνησε

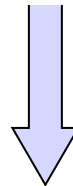
(Carlos Alchourron and David Makinson)

Παλιός Νομικός Κώδικας

- Τα παιδιά μπορούν να δουν TV μόνο αν έχουν φάει βραδινό.
- Τα παιδιά μπορούν να φάνε βραδινό μόνο αφού τελειώσουν το διάβασμα.



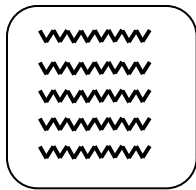
Τροπολογία: Τις Παρασκευές τα παιδιά μπορούν να δουν TV και χωρίς να έχουν διαβάσει.



Νέος Νομικός Κώδικας



2^η Αφιετηρία: *Semantics for Counterfactuals* (Peter Gardernfors)



⊨

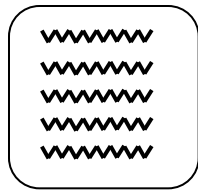
*Εάν ο Ναπολέων δεν είχε εισβάλει στην Ρωσία,
τότε θα είχε κατακτήσει ολόκληρη την Ευρώπη.*

K

A	B	A → B
T	T	T
T	F	F
F	T	T
F	F	T

2^η Αφιετηρία: Semantics for Counterfactuals

(Peter Gardernfors)



\models

Εάν ο Ναπολέων δεν είχε εισβάλει στην Ρωσία,
τότε θα είχε κατακτήσει ολόκληρη την Ευρώπη.

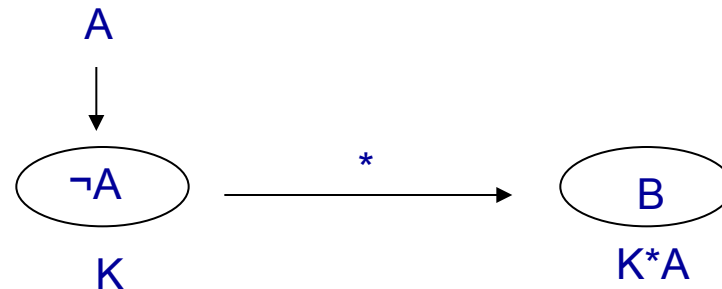
K

A	B	$A \rightarrow B$
T	T	T
T	F	F
F	T	T
F	F	T

Ramsey Test:

Η πρόταση "Εάν A , τότε B " αληθεύει στην γνωστική κατάσταση K αν το B αληθεύει στην γνωστική κατάσταση που προκύπτει από την αναθεώρηση του K ως προς A .

$$K \models A > B \quad \text{iff} \quad (K * A) \models B$$



Η Γέννηση του Belief Revision (1985)

C. Alchourron, P. Gardernfors, and D. Makinson, “On the logic of theory change: Partial meet functions for contraction and revision”, *Journal of Symbolic Logic*, 1985



**The AGM Framework
for
Theory Change**

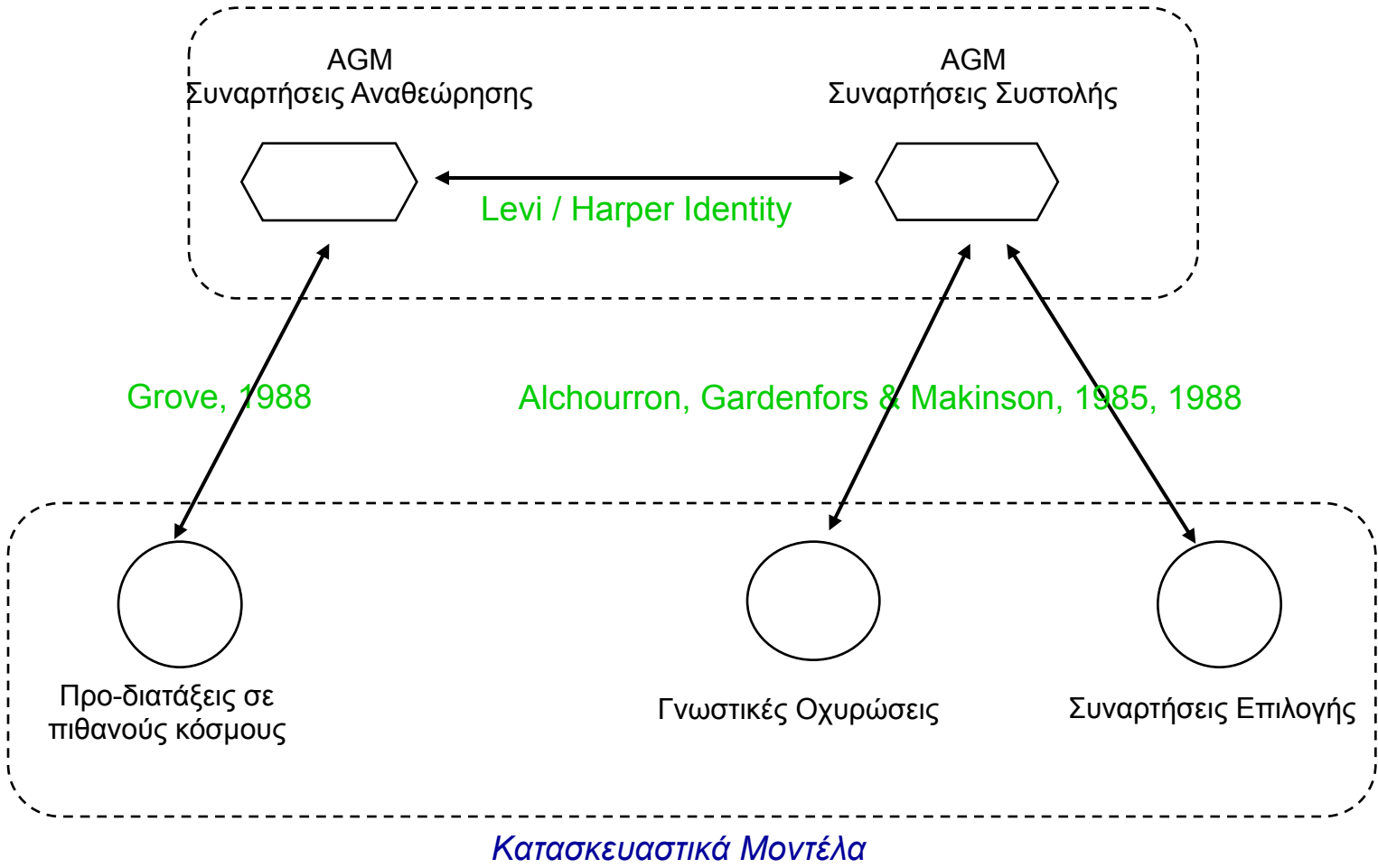
Χρήση Λογικής στην Μοντελοποίηση της Αναθεώρησης Πεποιθήσεων



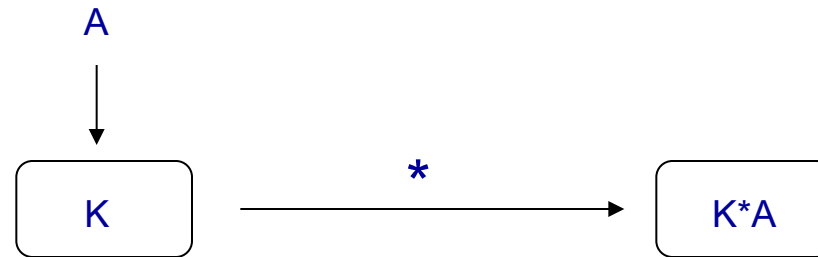
- Οι πεποιθήσεις αναπαριστώνται ως τύποι Προτασιακής Λογικής.
- Οι γνωστικές καταστάσεις αναπαριστώνται ως σύνολα τύπων, κλειστών ως προς την λογική συνεπαγωγή (αν $T \models \phi$, τότε $\phi \in T$).

Μοντέλα για Αναθεώρηση Πεποιθήσεων

Αξιωματικά Μοντέλα

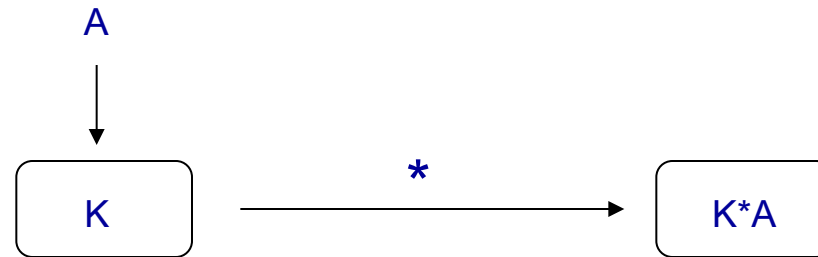


Τα Αξιώματα AGM για Αναθεώρηση Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής: Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της νέας πληροφορίας.

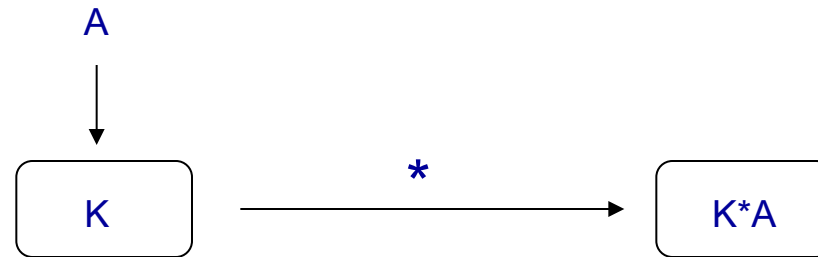
Τα Αξιώματα AGM για Αναθεώρηση Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής: Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της νέας πληροφορίας.

(K*1) Το $K*A$ είναι θεωρία

Τα Αξιώματα AGM για Αναθεώρηση Πεποιθήσεων

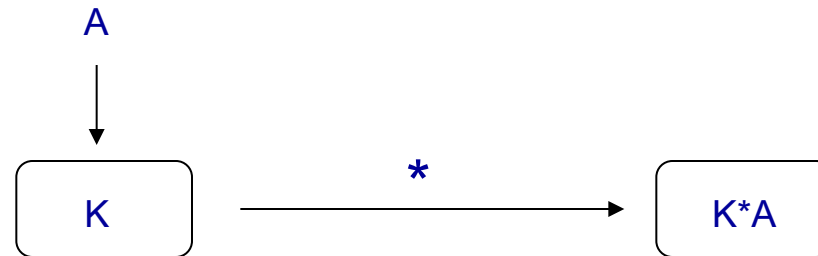


Αρχή της Ελάχιστης Αλλαγής: Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της νέας πληροφορίας.

(K*1) Το $K*A$ είναι θεωρία

(K*2) $A \in (K*A)$

Τα Αξιώματα AGM για Αναθεώρηση Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής: Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της νέας πληροφορίας.

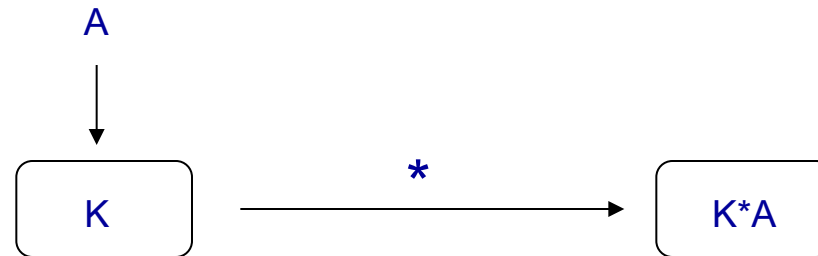
(K*1) Το $K*A$ είναι θεωρία

(K*2) $A \in (K*A)$

(K*3) $K*A \subseteq K+A$

(K*4) $\text{Αν } \neg A \notin K \text{ τότε } K+A \subseteq K*A$

Τα Αξιώματα AGM για Αναθεώρηση Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής: Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της νέας πληροφορίας.

(K*1) Το $K*A$ είναι θεωρία

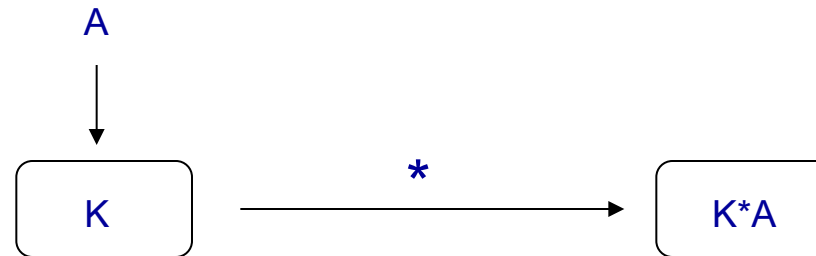
(K*2) $A \in (K*A)$

(K*3) $K*A \subseteq K+A$

(K*4) Αν $\neg A \notin K$ τότε $K+A \subseteq K*A$

(K*5) Αν το A δεν είναι αντίφαση, τότε το $K*A$ δεν περιέχει αντιφάσεις

Τα Αξιώματα AGM για Αναθεώρηση Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής: Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της νέας πληροφορίας.

(K*1) Το $K*A$ είναι θεωρία

(K*2) $A \in (K*A)$

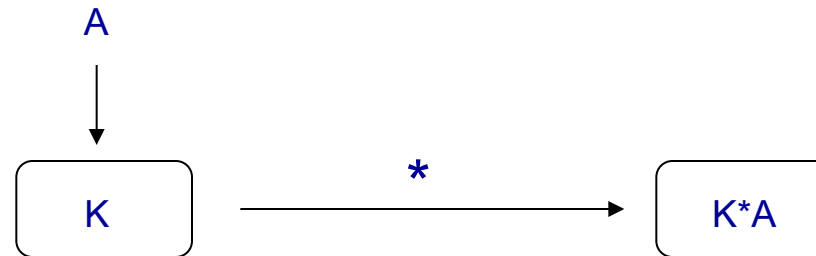
(K*3) $K*A \subseteq K+A$

(K*4) Αν $\neg A \notin K$ τότε $K+A \subseteq K*A$

(K*5) Αν το A δεν είναι αντίφαση, τότε το $K*A$ δεν περιέχει αντιφάσεις

(K*6) Αν $A \equiv B$ τότε $K*A = K*B$

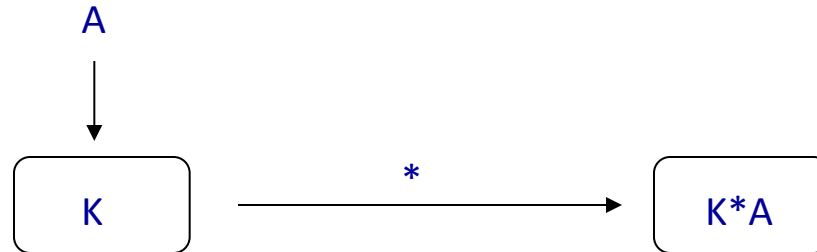
Τα Αξιώματα AGM για Αναθεώρηση Πεποιθήσεων



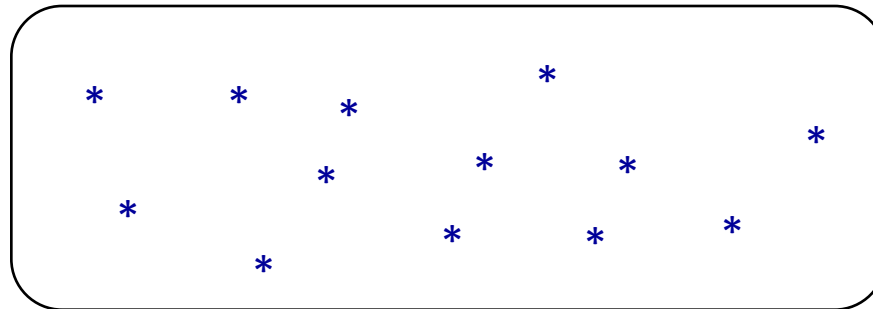
Αρχή της Ελάχιστης Αλλαγής: Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της νέας πληροφορίας.

- (K*1) Το K^*A είναι θεωρία
- (K*2) $A \in (K^*A)$
- (K*3) $K^*A \subseteq K+A$
- (K*4) Αν $\neg A \notin K$ τότε $K+A \subseteq K^*A$
- (K*5) Αν το A δεν είναι αντίφαση, τότε το K^*A δεν περιέχει αντιφάσεις
- (K*6) Αν $A \equiv B$ τότε $K^*A = K^*B$
- (K*7) $K^*(A \wedge B) \subseteq (K^*A)+B$
- (K*8) Αν $\neg B \notin (K^*A)$ τότε $(K^*A)+B \subseteq K^*(A \wedge B)$

Το Πλήθος των AGM Συναρτήσεων Αναθεώρησης

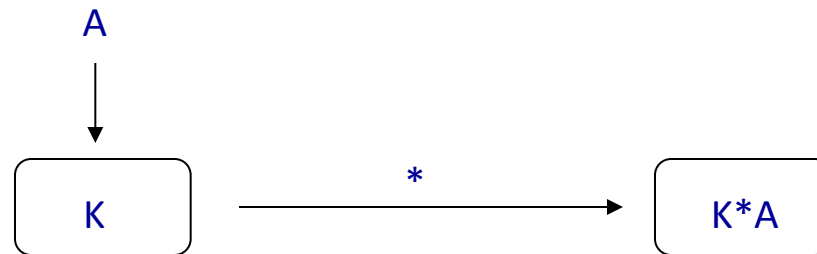


Δεδομένης μιας αρχικής γνωστικής κατάστασης K και νέας πληροφορίας A , τα αξιώματα AGM δεν προσδιορίζουν μοναδιαία την νέα γνωστική κατάσταση $K*A$.



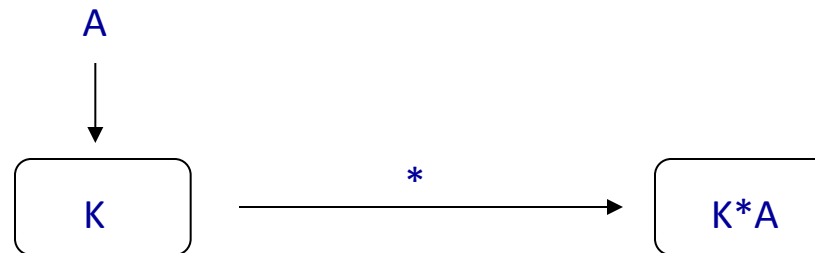
Συναρτήσεις που ικανοποιούν τα αξιώματα
($K*1$) - ($K*8$)

Ορισμένα Επιπλέον Αξιώματα



($K*M$) Αν $K \subseteq H$ τότε $K*A \subseteq H*A$.

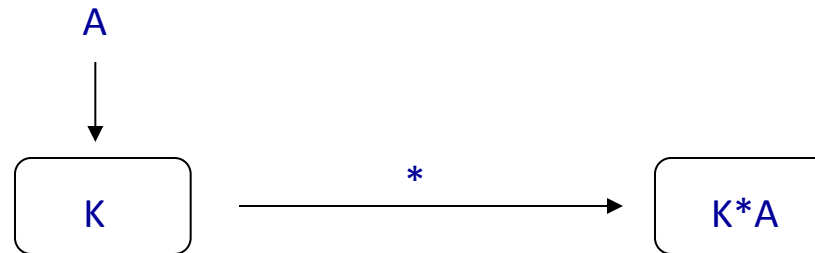
Ορισμένα Επιπλέον Αξιώματα



$(K*M)$ Αν $K \subseteq H$ τότε $K*A \subseteq H*A$.

Θεώρημα: Η συνθήκη $(K*M)$ έρχεται σε αντίθεση με τα $(K*1) - (K*8)$.

Ορισμένα Επιπλέον Αξιώματα

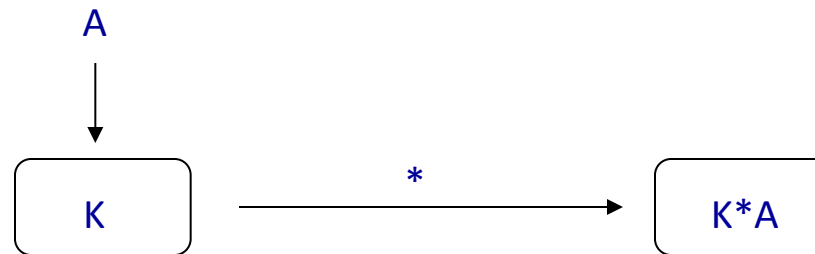


(K^*M) Αν $K \subseteq H$ τότε $K^*A \subseteq H^*A$.

Θεώρημα: Η συνθήκη (K^*M) έρχεται σε αντίθεση με τα $(K^*1) - (K^*8)$.

(K^*R) Αν $B \in K$ και $B \notin K^*A$ τότε $\neg B \in K^*A$.

Ορισμένα Επιπλέον Αξιώματα



(K^*M) Αν $K \subseteq H$ τότε $K^*A \subseteq H^*A$.

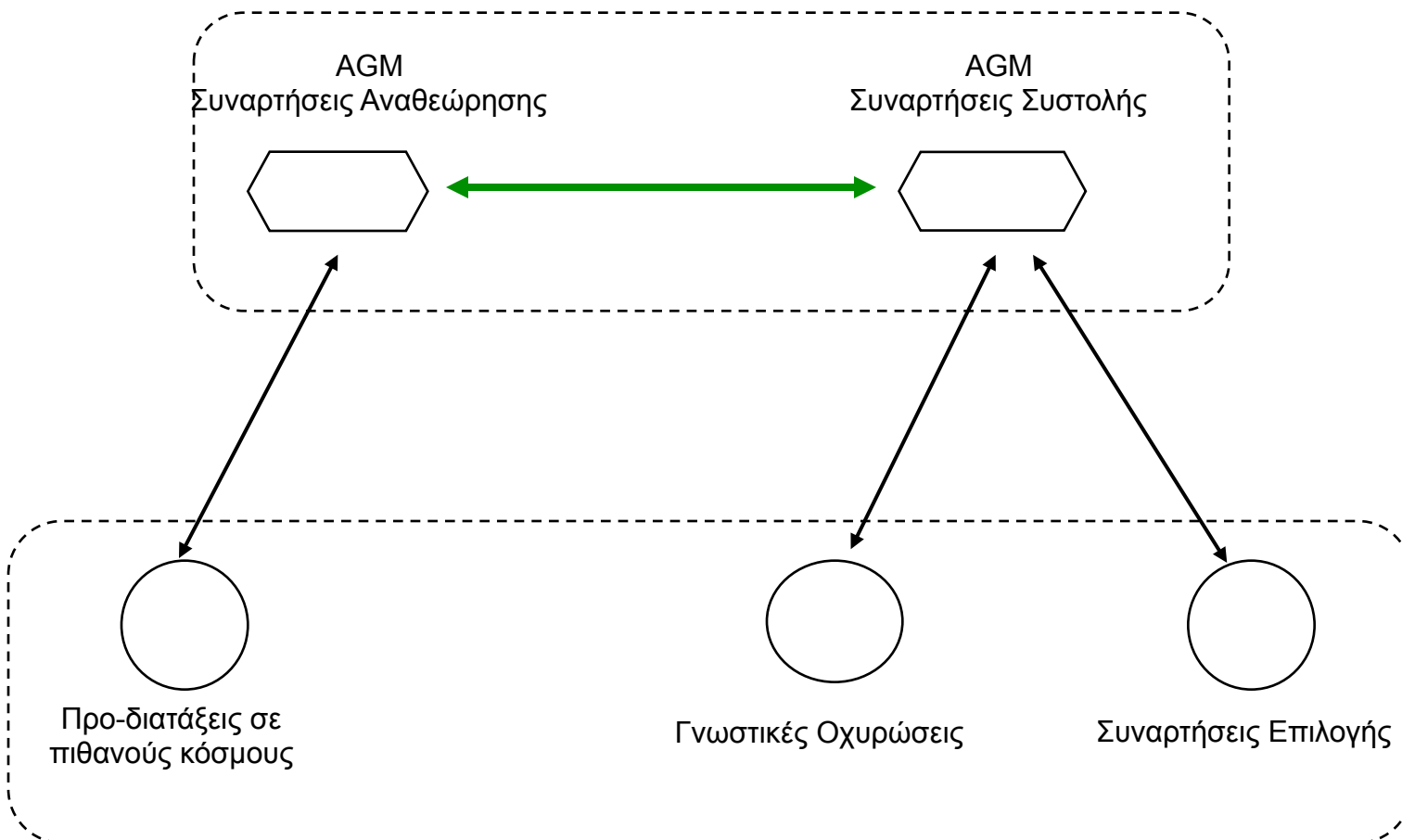
Θεώρημα: Η συνθήκη (K^*M) έρχεται σε αντίθεση με τα $(K^*1) - (K^*8)$.

(K^*R) Αν $B \in K$ και $B \notin K^*A$ τότε $\neg B \in K^*A$.

Θεώρημα: Εάν η $*$ ικανοποιεί τα $(K^*1) - (K^*8)$ και (K^*R) , τότε εφόσον $\neg A \in K$ το K^*A είναι πλήρες.

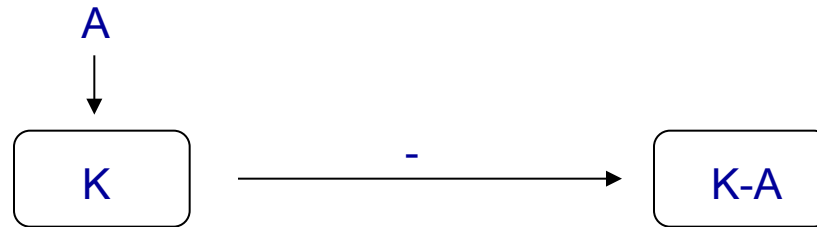
Μοντέλα για Αναθεώρηση Πεποιθήσεων

Αξιωματικά Μοντέλα



Κατασκευαστικά Μοντέλα

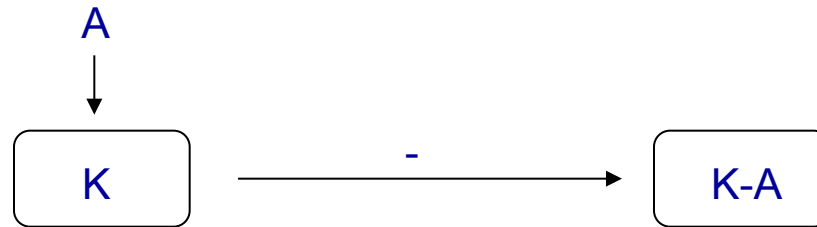
Τα Αξιώματα AGM για Συστολή Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής:

Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της A .

Τα Αξιώματα AGM για Συστολή Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής:

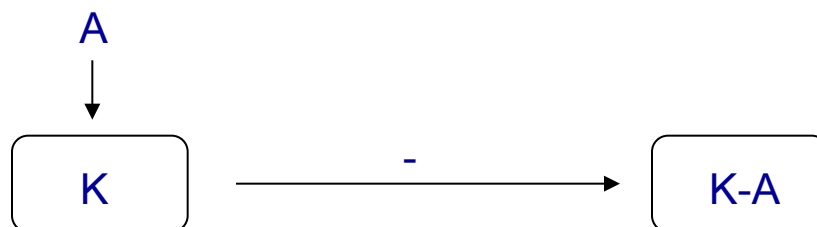
Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της A .

(K-1) Το $K-A$ είναι θεωρία

(K-2) $K-A \subseteq K$

(K-3) Αν $A \notin K$ τότε $K - A = K$

Τα Αξιώματα AGM για Συστολή Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής:

Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της A .

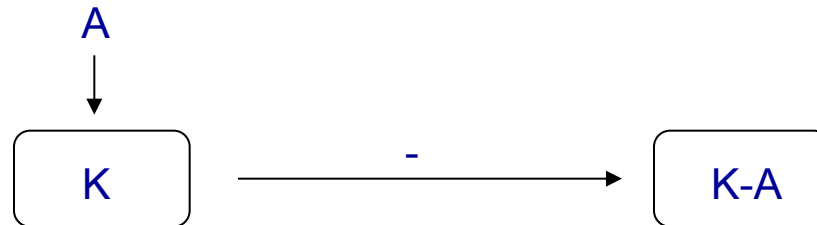
(K-1) Το $K-A$ είναι θεωρία

(K-2) $K-A \subseteq K$

(K-3) Αν $A \notin K$ τότε $K - A = K$

(K-4) Αν $\neq A$ τότε $A \notin K-A$

Τα Αξιώματα AGM για Συστολή Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής:

Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της A .

(K-1) Το $K-A$ είναι θεωρία

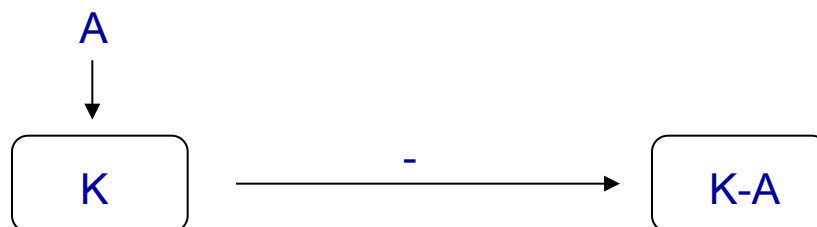
(K-2) $K-A \subseteq K$

(K-3) Αν $A \notin K$ τότε $K - A = K$

(K-4) Αν $\neq A$ τότε $A \notin K-A$

(K-5) Αν $A \in K$ τότε $K \subseteq (K-A)+A$

Τα Αξιώματα AGM για Συστολή Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής:

Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της A .

(K-1) Το $K-A$ είναι θεωρία

(K-2) $K-A \subseteq K$

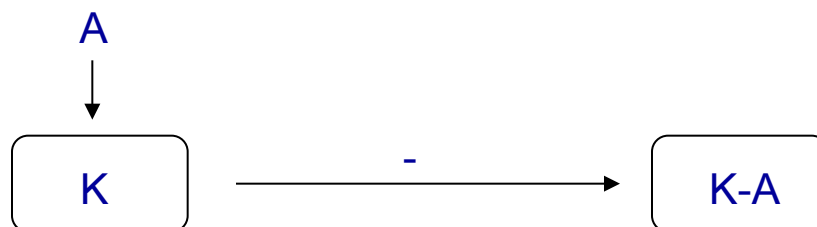
(K-3) Αν $A \notin K$ τότε $K - A = K$

(K-4) Αν $\neq A$ τότε $A \notin K-A$

(K-5) Αν $A \in K$ τότε $K \subseteq (K-A)+A$

(K-6) Αν $A \equiv B$ τότε $K-A = K-B$

Τα Αξιώματα AGM για Συστολή Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής:

Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της A .

(K-1) Το $K-A$ είναι θεωρία

(K-2) $K-A \subseteq K$

(K-3) Αν $A \notin K$ τότε $K - A = K$

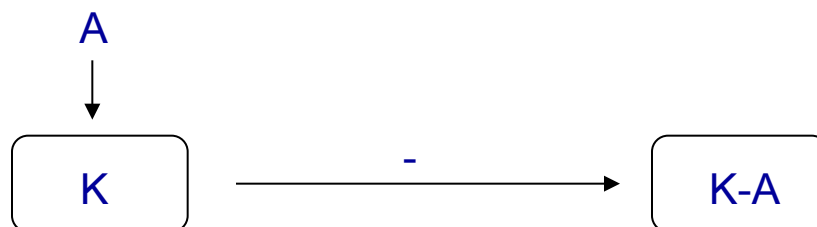
(K-4) Αν $\neg A$ τότε $A \notin K-A$

(K-5) Αν $A \in K$ τότε $K \subseteq (K-A)+A$

(K-6) Αν $A \equiv B$ τότε $K-A = K-B$

(K-7) $(K-A) \cap (K-B) \subseteq K-(A \wedge B)$

Τα Αξιώματα AGM για Συστολή Πεποιθήσεων



Αρχή της Ελάχιστης Αλλαγής:

Η νέα γνωστική κατάσταση θα πρέπει να διαφέρει όσο γίνεται λιγότερο από την αρχική, δεδομένης της A .

(K-1) Το $K-A$ είναι θεωρία

(K-2) $K-A \subseteq K$

(K-3) Αν $A \notin K$ τότε $K - A = K$

(K-4) Αν $\models A$ τότε $A \notin K-A$

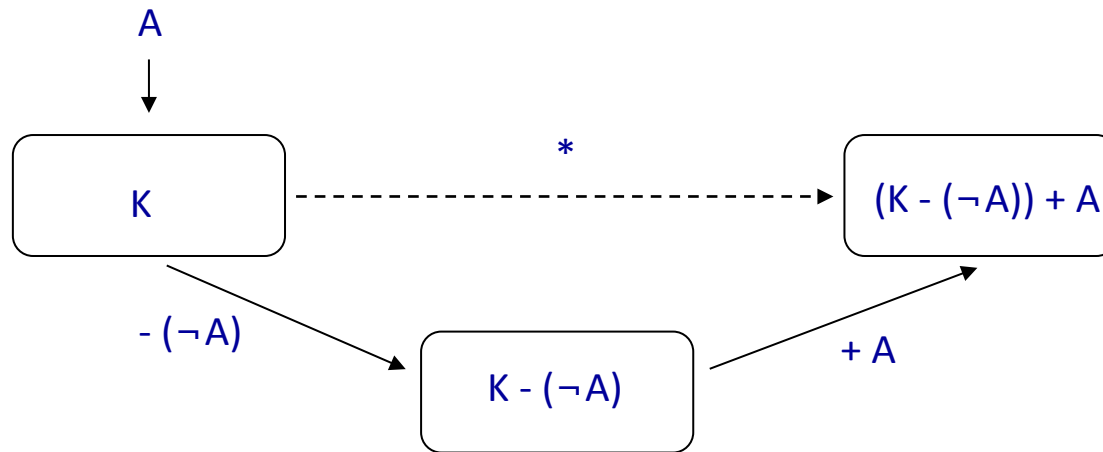
(K-5) Αν $A \in K$ τότε $K \subseteq (K-A)+A$

(K-6) Αν $A \equiv B$ τότε $K-A = K-B$

(K-7) $(K-A) \cap (K-B) \subseteq K-(A \wedge B)$

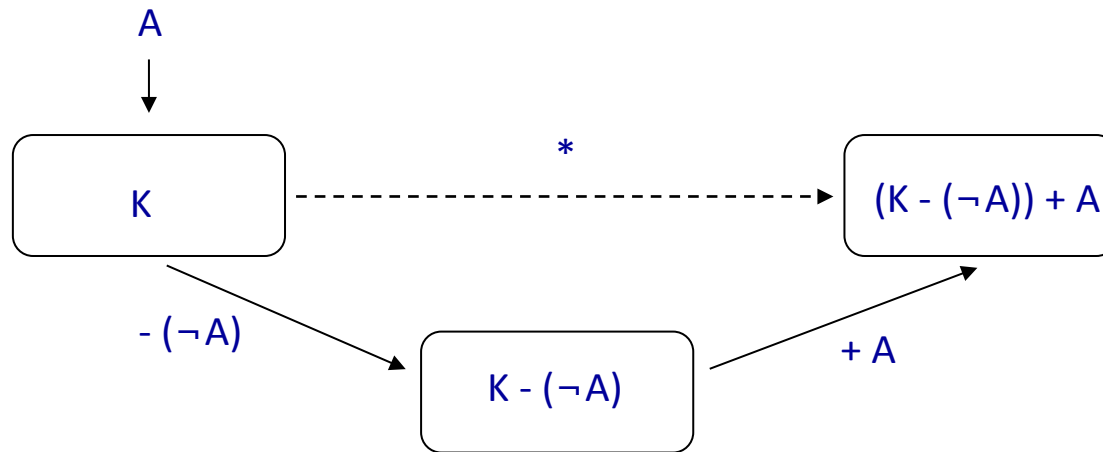
(K-8) Αν $A \notin K-(A \wedge B)$ τότε $K-(A \wedge B) \subseteq K-A$

Levi Identity



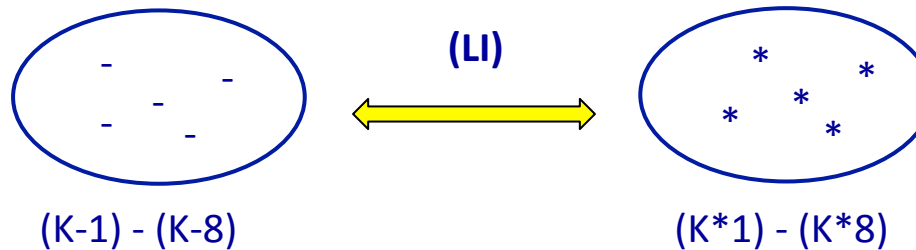
$$(LI) \quad K * A = (K - (\neg A)) + A$$

Levi Identity



$$(LI) \quad K * A = (K - (\neg A)) + A$$

Θεώρημα:



Levi Identity – Απόδειξη

Βοηθητικοί Ορισμοί/Αποτελέσματα

Ορισμός: $[T] = \{ M : M \models T \}$

Ορισμός: $T+H = \{ \phi \in L : M \models \phi, \text{ για κάθε } M \in [T] \cap [H] \}$

Χρήσιμα Λήμματα:

- $[T+H] = [T] \cap [H]$.
- Το $T+H$ είναι θεωρία.
- Αν $H \subseteq T$ τότε $[T] \subseteq [H]$.
- $[T \cup H] = [T] \cap [H]$.
- Αν $\psi \in K+\phi$ τότε $\phi \Rightarrow \psi \in K$.
- $T+(A \wedge B) = (T+A)+B$.

Levi Identity – Απόδειξη

$$(LI) \quad K^*A = (K - (\neg A)) + A$$

(K*1) Το K^*A είναι θεωρία

(K*2) $A \in (K^*A)$

(K*3) $K^*A \subseteq K+A$

(K*4) Αν $\neg A \notin K$ τότε $K+A \subseteq K^*A$

(K*5) Αν το $A \neq \perp$, τότε $K^*A \neq \perp$

(K*6) Αν $A \equiv B$ τότε $K^*A = K^*B$

(K*7) $K^*(A \wedge B) \subseteq (K^*A)+B$

(K*8) Αν $\neg B \notin (K^*A)$ τότε $(K^*A)+B \subseteq K^*(A \wedge B)$

(K-1) Το $K-A$ είναι θεωρία ✓

(K-2) $K-A \subseteq K$ ✓

(K-3) Αν $A \notin K$ τότε $K - A = K$ ✓

(K-4) Αν $\neq A$ τότε $A \notin K-A$ ✓

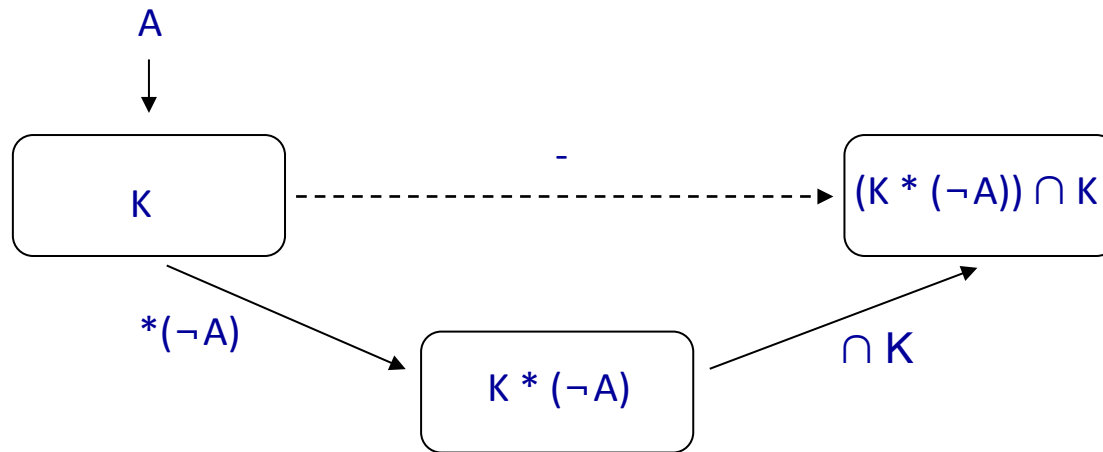
(K-5) Αν $A \in K$ τότε $K \subseteq (K-A)+A$ ✓

(K-6) Αν $A \equiv B$ τότε $K-A = K-B$ ✓

(K-7) $(K-A) \cap (K-B) \subseteq K-(A \wedge B)$ ✓

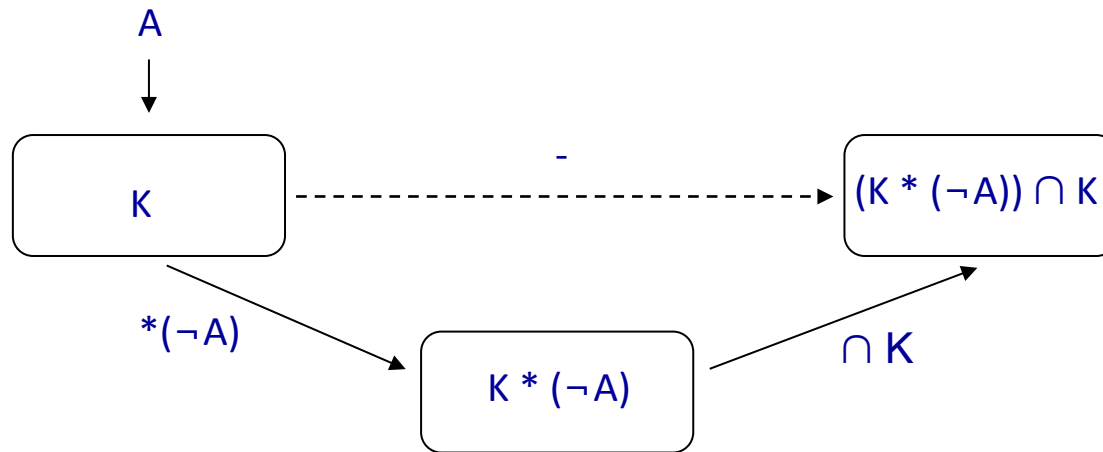
(K-8) Αν $A \notin K-(A \wedge B)$ τότε $K-(A \wedge B) \subseteq K-A$ ✓

Harper Identity



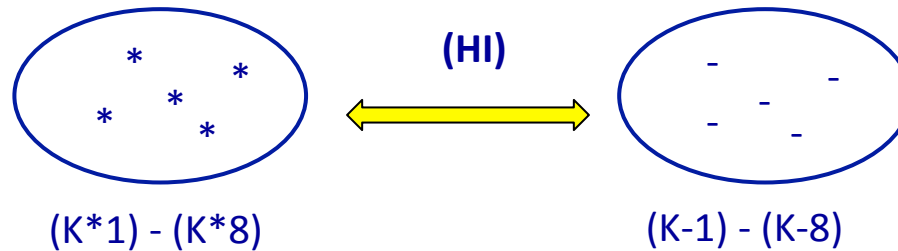
$$(HI) \quad K-A = (K * (\neg A)) \cap K$$

Harper Identity

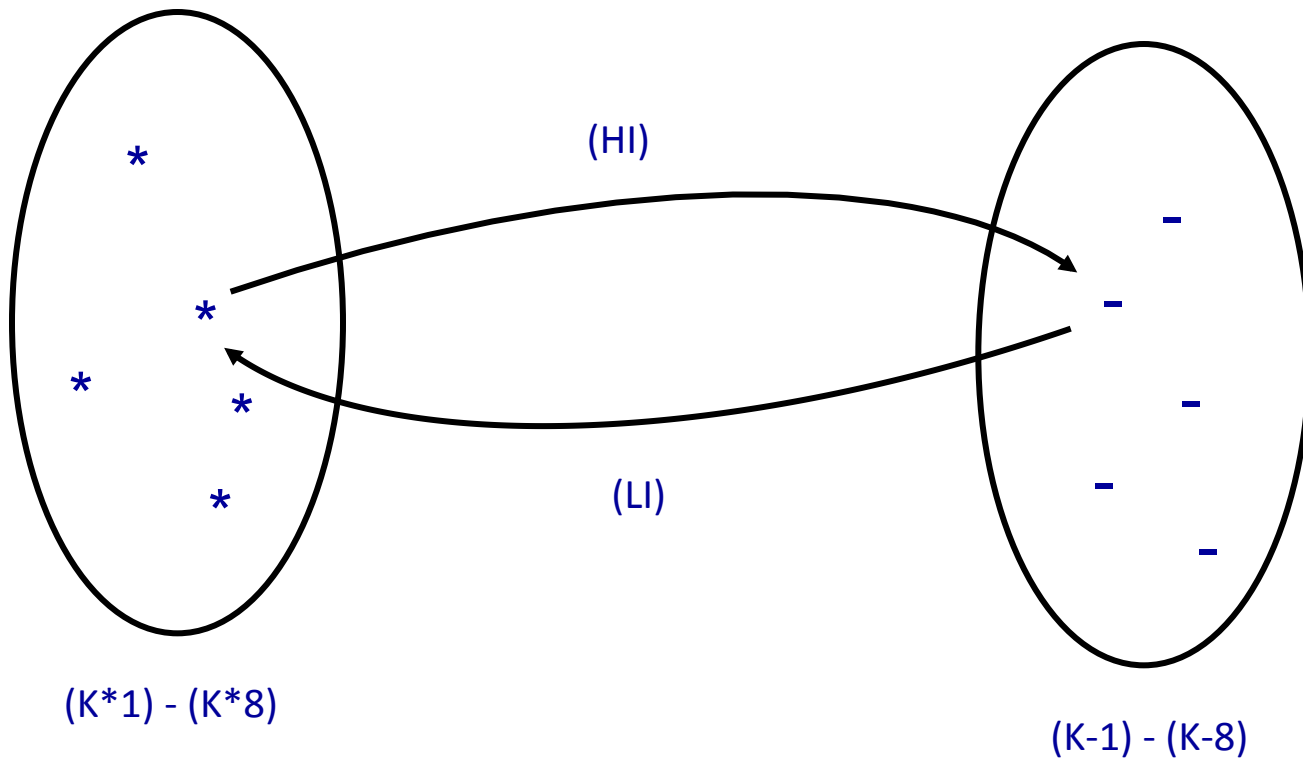


$$(HI) \quad K-A = (K * (\neg A)) \cap K$$

Θεώρημα:

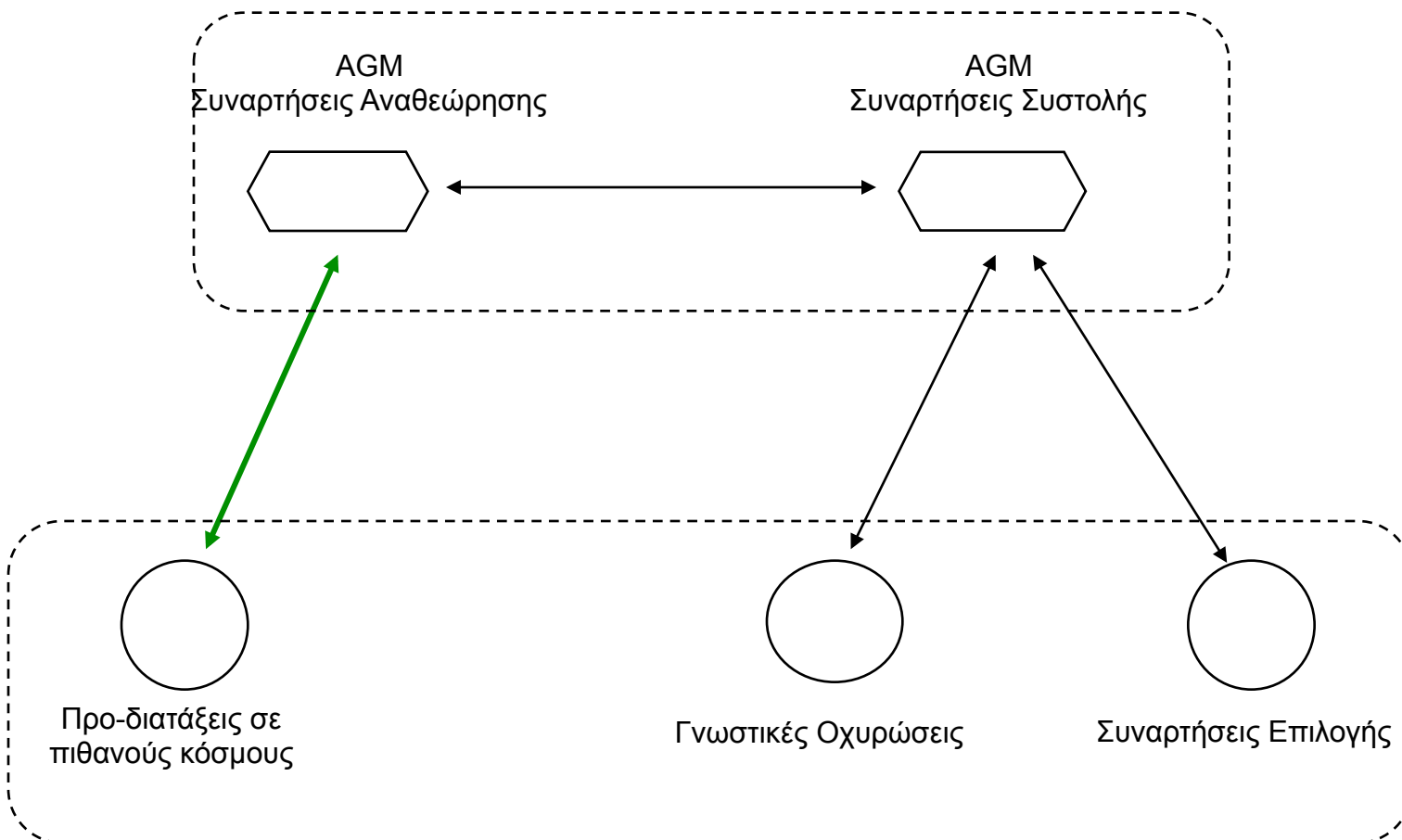


Inter-definability



Μοντέλα για Αναθεώρηση Πεποιθήσεων

Αξιωματικά Μοντέλα



Κατασκευαστικά Μοντέλα

Ένας Άλλος Κόσμος

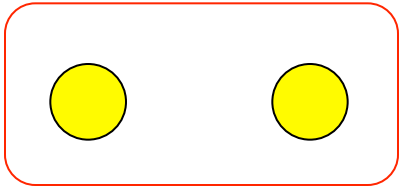
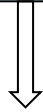


Πιθανοί Κόσμοι vs Προτάσεις

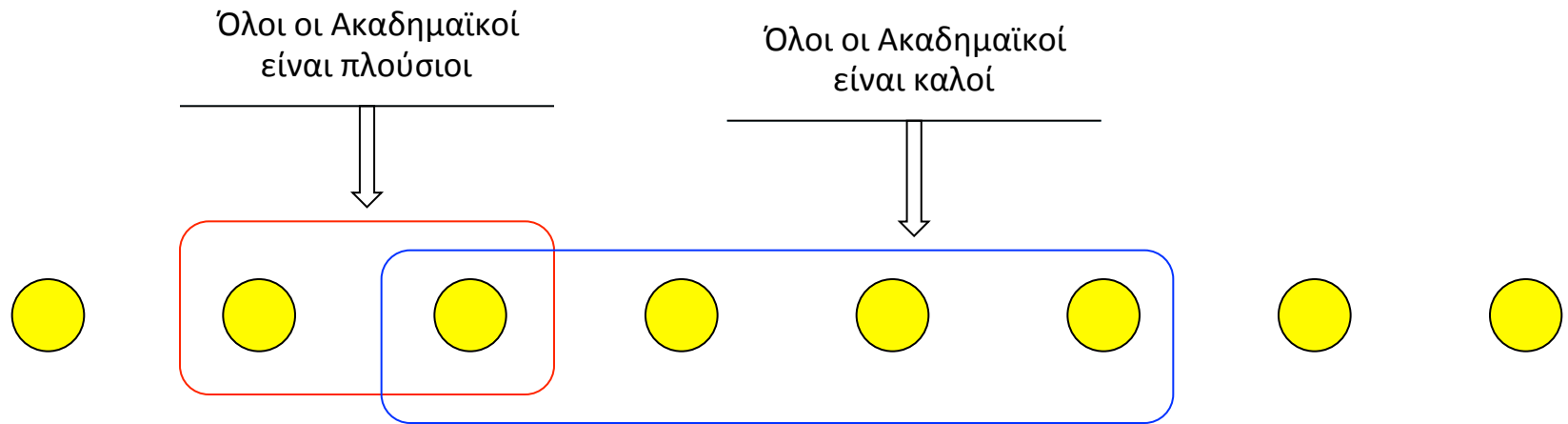


Πιθανοί Κόσμοι vs Προτάσεις

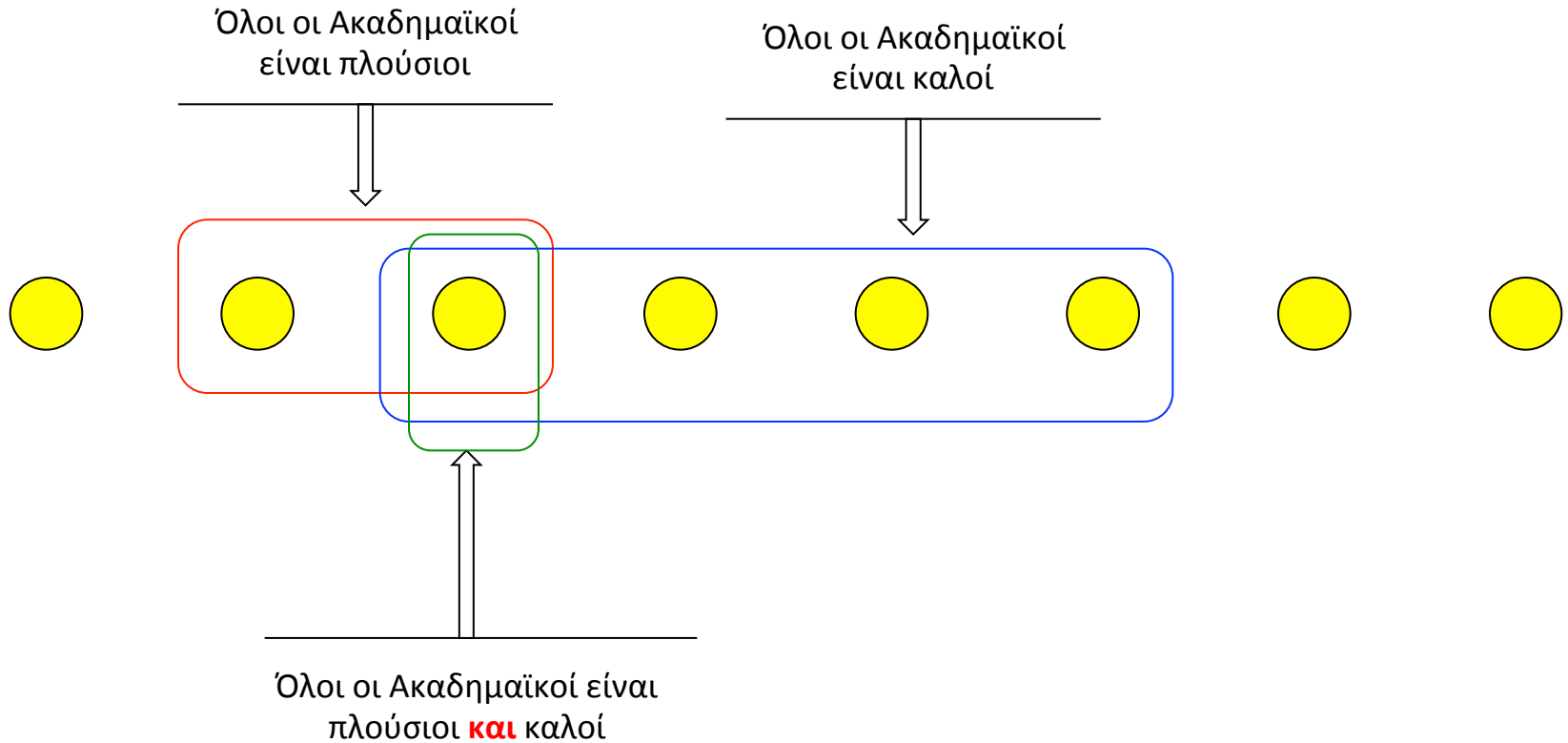
Όλοι οι Ακαδημαϊκοί
είναι πλούσιοι



Πιθανοί Κόσμοι vs Προτάσεις



Πιθανοί Κόσμοι vs Προτάσεις



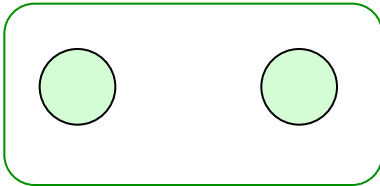
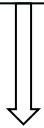
Αναθεώρηση Πεποιθήσεων μέσω Πιθανών Κόσμων

Το αμάξι του Γιάννη είναι BMM.

Τα BMM κατασκευάζονται στην Γαλλία

Η Γαλλία ανήκει στην ΕΕ.

Όλα αμάξια κατασκευάζονται στην ΕΕ
παίρνουν αμόλυβδη.



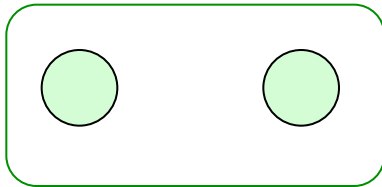
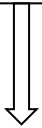
Αναθεώρηση Πεποιθήσεων μέσω Πιθανών Κόσμων

Το αμάξι του Γιάννη είναι BMM.

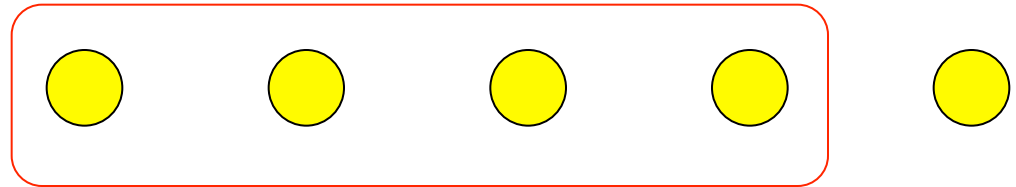
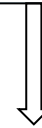
Τα BMM κατασκευάζονται στην Γαλλία

Η Γαλλία ανήκει στην ΕΕ.

Όλα αμάξια κατασκευάζονται στην ΕΕ
παίρνουν αμόλυβδη.



Το αμάξι του Γιάννη δεν παίρνει αμόλυβδη.



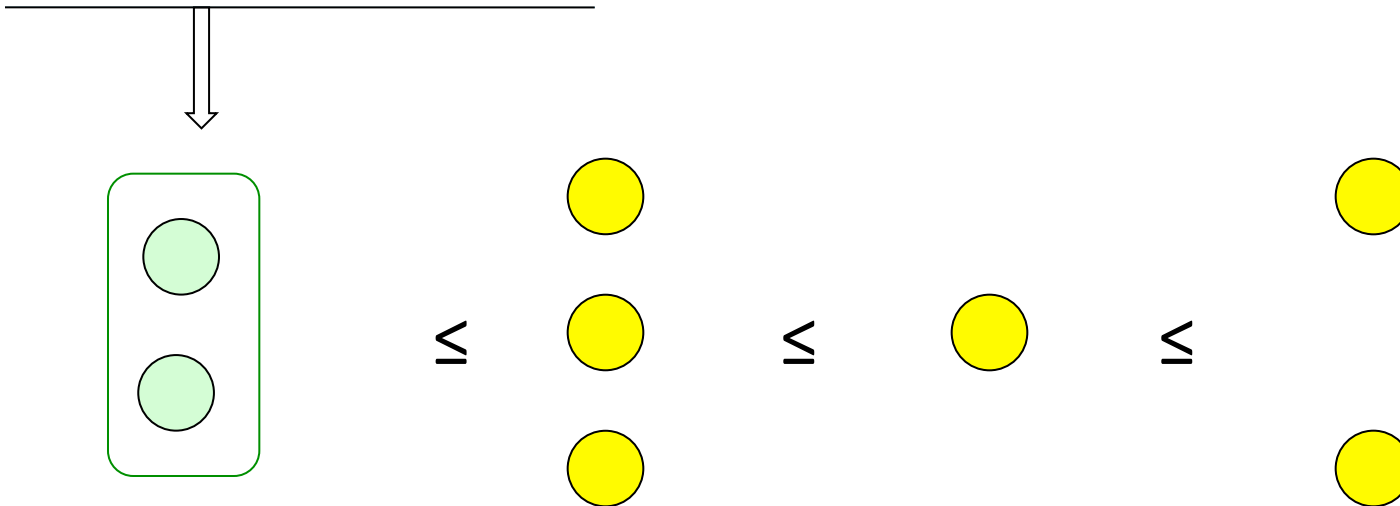
Διάταξη Ευλογοφάνειας

Το αμάξι του Γιάννη είναι BMM.

Τα BMM κατασκευάζονται στην Γαλλία

Η Γαλλία ανήκει στην ΕΕ.

Όλα αμάξια κατασκευάζονται στην ΕΕ
παίρνουν αμόλυβδη.



Διάταξη Ευλογοφάνειας

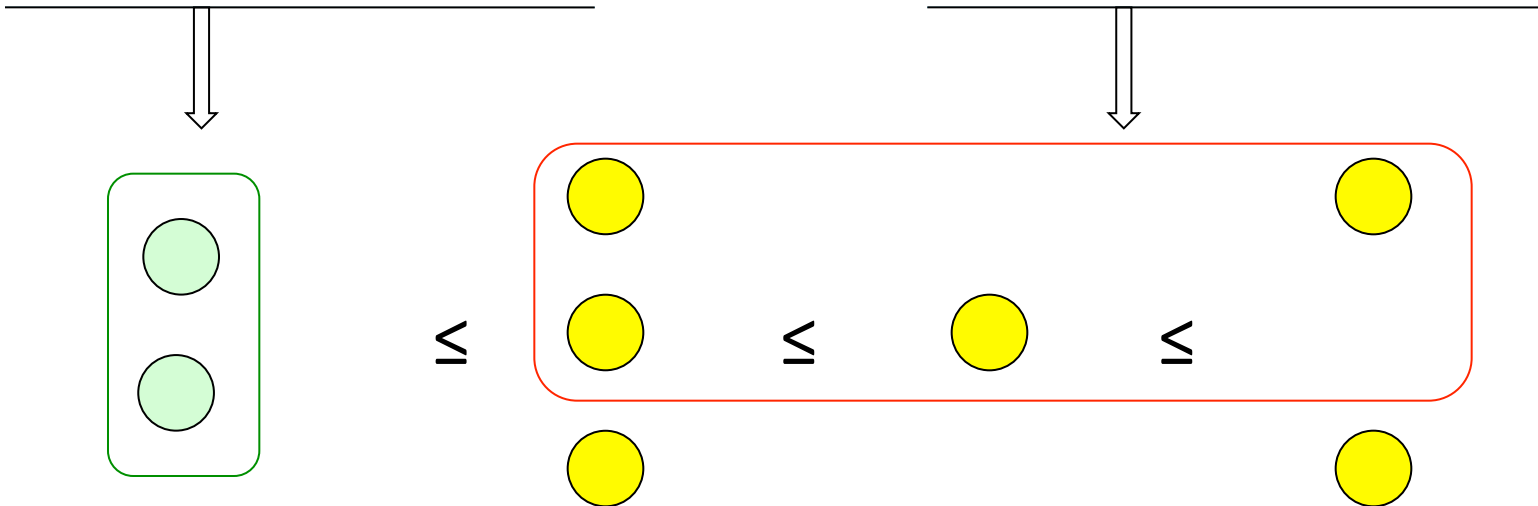
Το αμάξι του Γιάννη είναι BMM.

Τα BMM κατασκευάζονται στην Γαλλία

Η Γαλλία ανήκει στην ΕΕ.

Όλα αμάξια κατασκευάζονται στην ΕΕ
παίρνουν αμόλυβδη.

Το αμάξι του Γιάννη δεν παίρνει αμόλυβδη.



Διάταξη Ευλογοφάνειας

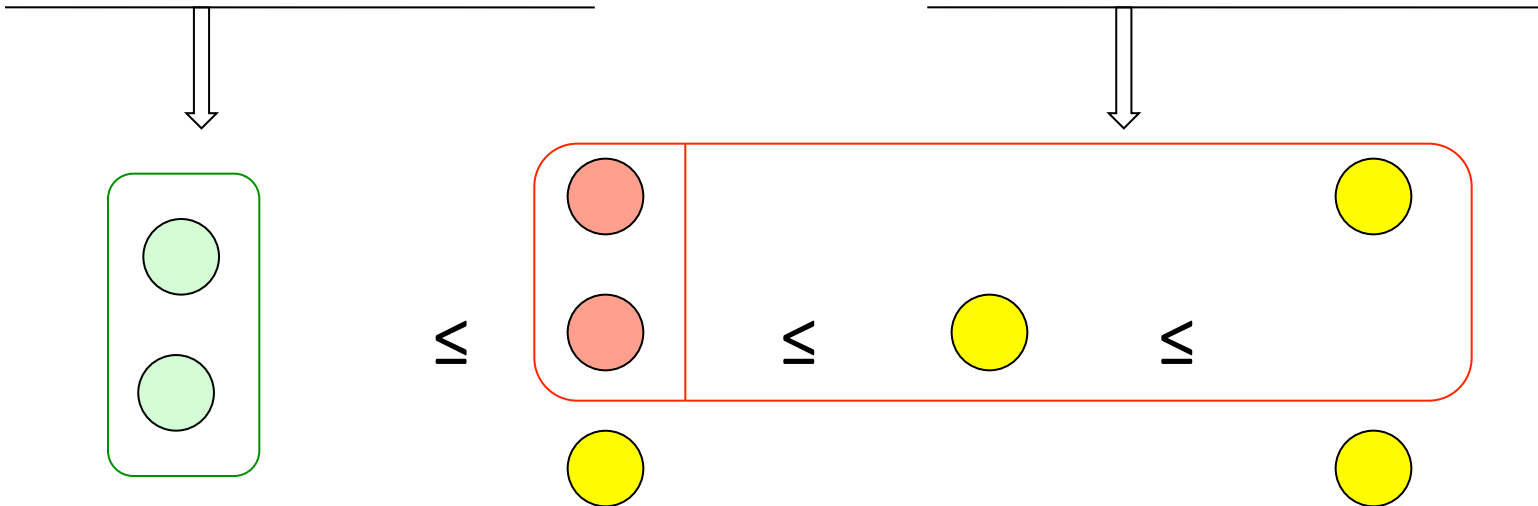
Το αμάξι του Γιάννη είναι BMM.

Τα BMM κατασκευάζονται στην Γαλλία

Η Γαλλία ανήκει στην ΕΕ.

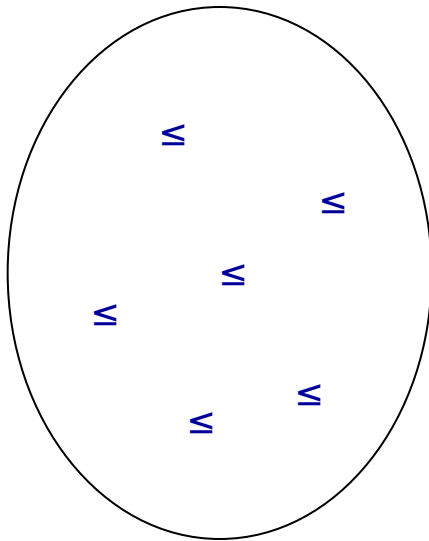
Όλα αμάξια κατασκευάζονται στην ΕΕ
παίρνουν αμόλυβδη.

Το αμάξι του Γιάννη δεν παίρνει αμόλυβδη.

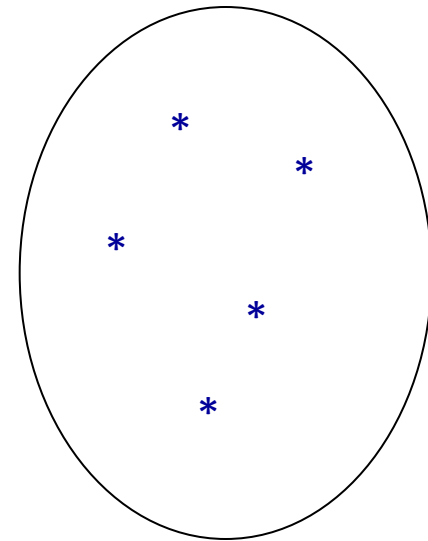
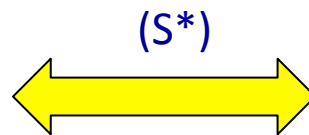


$$(S^*) \quad [K^*A] = \min([A], \leq)$$

Θεώρημα Αναπαράστασης



Προ-διατάξεις σε
Πιθανούς Κόσμους



Συναρτήσεις Αναθεώρησης
 $(K^*1) - (K^*8)$