Τυπικές αποδείξεις (formal deductions) για το δίλημμα των τριών φυλακισμένων

Χρησιμοποιούμε το γενικό μοντέλο **P**: οι καταστάσεις είναι *όλες* οι δυνατές αναθέσεις χρωμάτων στους τρείς παίκτες.

Μία φόρμουλα φ *ισχύει* σε ένα μοντέλο Kripke, όταν η φ αληθεύει *σε κάθε κατάσταση*.

I Οι φόρμουλες

$$\neg(K_A \text{ AisWh}) \rightarrow (\text{BisWh} \vee \text{CisWh})$$
$$\neg(K_B \text{ AisWh}) \rightarrow (\text{AisWh} \vee \text{CisWh})$$
$$\neg(K_C \text{ AisWh}) \rightarrow (\text{BisWh} \vee \text{AisWh})$$

ισχύουν στο P.

Άρα, θα ισχύουν οι εξής clauses στο P:

<u>clause 1</u>: $(K_A \text{ AisWh}) \vee \text{BisWh} \vee \text{CisWh}$

<u>clause 2</u>: $(K_B \text{ BisWh}) \vee \text{AisWh} \vee \text{CisWh}$

<u>clause 3</u>: $(K_C \text{ CisWh}) \vee \text{BisWh} \vee \text{AisWh}$

Παρατήρηση 1 Οι clauses 1 , 2 , 3 θα ισχύουν σε οποιοδήποτε μοντέλο που αποτελείται από ένα *υποσύνολο* των καταστάσεων του P,
και τα possibility relations ορίζονται όπως στο P .

II Η φόρμουλα $\neg(K_A \text{ AisWh})$ αληθεύει στην κατάσταση u , όταν:

u(AisWh) = *false* ,

ή

u(AisWh) = *true* , και

υπάρχει (μία τουλάχιστον) κατάσταση v , όπου

u AisWh

A v $\neg$AisWh

1. u $\approx_A$ v άρα v(BisWh) = u(BisWh) ,

v(CisWh) = u(CisWh)

2. v(AisWh) = *false* .

Αντίστοιχα για τις φόρμουλες $\neg(K_B \text{ BisWh})$, $\neg(K_C \text{ CisWh})$.

Παρατήρηση 2 Το ***II*** εφαρμόζεται σε οποιοδήποτε μοντέλο όπου τα possibility relations ορίζονται όπως στο P .

Ανάλυση για την **Case 3**

Ισχύουν οι εξής clauses στο P:

clause 1: $(K_A \text{ AisWh}) \vee \text{BisWh} \vee \text{CisWh}$

clause 2: $(K_B \text{ BisWh}) \vee \text{AisWh} \vee \text{CisWh}$

clause 3: $(K_C \text{ CisWh}) \vee \text{BisWh} \vee \text{AisWh}$

***Αμέσως μετά το* Step 3.1:**

Παίρνοντας υπόψη τις ανακοινώσεις των παικτών, εξετάζουμε *τις clauses*

clause 4: $\neg(K_A \text{ AisWh})$ clause 5: $\neg(K_B \text{ BisWh})$ clause 6: $\neg(K_C \text{ CisWh})$

Θεώρημα 1 $\{ \text{clause1}, \text{clause2}, \text{clause3} \} \cup \{ \text{clause4}, \text{clause5}, \text{clause6} \}$

$\models (\text{BisWh} \vee \text{CisWh}) \wedge (\text{AisWh} \vee \text{CisWh}) \wedge (\text{BisWh} \vee \text{AisWh})$

Απόδειξη Εφαρμόζουμε τον κανόνα της επίλυσης:

clause 1 with clause 4 gives clause 7: $\text{BisWh} \vee \text{CisWh}$

clause 2 with clause 5 gives clause 8: $\text{AisWh} \vee \text{CisWh}$

clause 3 with clause 6 gives clause 9: $\text{BisWh} \vee \text{AisWh}$ □

Εξετάζουμε *το σύνολο των καταστάσεων του* P (τις ονομάζουμε ενδεχόμενες καταστάσεις) *για τις οποίες αληθεύουν οι clauses* 4, 5, 6 .

Από την ***Παρατήρηση 1*** και το ***Θεώρημα 1*** : οι clauses 7, 8, 9 θα αληθεύουν σε κάθε κατάσταση του P που είναι ενδεχόμενη κατάσταση.

Παρατήρηση 3 Οι clauses 7, 8, 9 αληθεύουν.σε μία κατάσταση, *άν και μόνο άν* το πολύ ένας παίκτης έχει μαύρο χρώμα.

clause 7: $BisWh \vee CisWh$ clause 8: $AisWh \vee CisWh$ clause 9: $BisWh \vee AisWh$

clause 4: $\neg(K_A\ AisWh)$ clause 5: $\neg(K_B\ BisWh)$ clause 6: $\neg(K_C\ CisWh)$

Θεώρημα 2 { $\forall$s ' clause7 $\wedge$ clause8 $\wedge$ clause9 αληθεύει στην s ' }

$\cup$ { clause4 , clause5 , clause6 }

|= (BisWh $\wedge$ CisWh $\wedge$ AisWh)

Απόδειξη

Έστω u μία κατάσταση όπου αληθεύουν οι υποθέσεις της συνεπαγωγής.

(1) Εξετάζουμε την clause 4, $\neg(K_A\ AisWh)$, η οποία αληθεύει στην u .
Θα δείξουμε ότι οι clauses CisWh , BisWh θα αληθεύουν στην u.

α Αν u(AisWh) = true :

Επειδή η φόρμουλα $\neg(K_A\ AisWh)$ αληθεύει στην u , θα υπάρχει (λόγω του ***II***)
μία κατάσταση v στην οποία θα αληθεύει η clause $\neg$AisWh .

u AisWh

A v $\neg$AisWh

Εφαρμόζουμε τον κανόνα της επίλυσης:

$\neg$AisWh with clause 8 gives CisWh

$\neg$AisWh with clause 9 gives BisWh

Άρα οι clauses CisWh , BisWh θα αληθεύουν στην v .

Λόγω της επιλογής της v θα έχουμε v(BisWh) = u(BisWh) , v(CisWh) = u(CisWh) .

Άρα οι clauses CisWh , BisWh θα αληθεύουν και στην u .

β Αν u(AisWh) = false : η clause $\neg$AisWh αληθεύει στην u .

Χρησιμοποιώντας τις clauses 8, 9 όπως στην προηγούμενη περίπτωση
βρίσκουμε ότι οι clauses CisWh , BisWh θα αληθεύουν στην u.

(2) Εξετάζουμε την clause 5, $\neg(K_B\ BisWh)$, η οποία αληθεύει στην u .
Με ανάλογο τρόπο, βρίσκουμε ότι και η clause AisWh θα αληθεύει στην u . □

Έστω **Q** ένα μοντέλο που αποτελείται από τις <u>ενδεχόμενες καταστάσεις του P</u>,
με possibility relations που ορίζονται όπως στο P.

1 Επειδή οι clauses 1, 2, 3, 7, 8, 9 αληθεύουν σε κάθε ενδεχόμενη κατάσταση του P, οι clauses 1, 2, 3, 7, 8, 9 *θα ισχύουν στο Q* (Παρατηρήσεις 1 και 3)..

2 Δεν είναι σαφές αν οι clauses 4, 5, 6 ισχύουν επίσης στο Q .

***Αμέσως μετά το* Step 3.2:**

Παίρνοντας υπόψη τις ανακοινώσεις των παικτών, εξετάζουμε *σε ποιές καταστάσεις του* Q αληθεύουν οι clauses 4 , 5 , 6.

<u>clause 4</u>: $\neg(K_A$ AisWh$)$ <u>clause 5</u>: $\neg(K_B$ BisWh$)$ <u>clause 6</u>: $\neg(K_C$ CisWh$)$

Από το ***Θεώρημα 2*** : η φόρμουλα BisWh $\wedge$ CisWh $\wedge$ AisWh θα αληθεύει
στις καταστάσεις του Q. όπου αληθεύουν οι clauses 4 , 5 , 6.

Προτεινόμενες ασκήσεις

1 Χρησιμοποιείστε φόρμουλες και τυπικές αποδείξεις για να αναλύσετε το πρόβλημα στην Exercise 1.3 - *Reasoning About Knowledge*.

2 Χρησιμοποιείστε φόρμουλες και τυπικές αποδείξεις για να αναλύσετε το Muddy Children Puzzle όταν $k = 2$ (δύο λερωμένα παιδάκια).

3 Χρησιμοποιείστε φόρμουλες και τυπικές αποδείξεις για να αναλύσετε την **Case 2.**

4 Συμπληρώστε την περίπτωση (2) της Απόδειξης του Θεωρήματος 2. Γιατί η clause AisWh θα αληθεύει στην κατάσταση u;