Άσκηση 1.3 από "Reasoning About Knowledge (1995)"

Καμινάρης Κωνσταντίνος

11/3/2024

Άσκηση 1.3

Ερώτηση

The wise men puzzle is a well-known variant of the muddy children puzzle. The standard version of the story goes as follows: There are three wise men. It is common knowledge that there are three red hats and two white hats. The king puts a hat on the head of each of the three wise men, and asks them (sequentially) if they know the color of the hat on their head. The first wise man says that he does not know; the second wise man says that he does not know; then the third wise man says that he knows.

α' What color is the third wise man's hat?

β' We have implicitly assumed in the story that the wise men can all see. Suppose we assume instead that the third wise man is blind and that it is common knowledge that the first two wise men can see. Can the third wise man still figure out the color of his hat?

Απάντηση για ταυτόχρονες αποκρίσεις

Στην ανάλυση που ακολουθεί έχει γίνει η σημαντική παραδοχή πως οι απαντήσεις των παικτών είναι ταυτόχρονες (και όχι με τη σειρά). Επίσης έχει γίνει η (λιγότερο σημαντική) παραδοχή πως τα καπέλα είναι 3 λευκά και 2 μαύρα αντί για τα 3 κόκκινα και 2 λευκά της εκφώνησης. Θα χρησιμοποιηθούν οι ατομικές προτάσεις $AisWh$, $BisWh$ και $CisWh$, με την ευκόλως εννοούμενη ερμηνεία τους.

	A	B	C
s1	W	W	W
s2	B	W	W
s3	W	B	W
s4	W	W	B
s5	B	B	W
s6	B	W	B
s7	W	B	B

Table 1: Πιθανοί κόσμοι

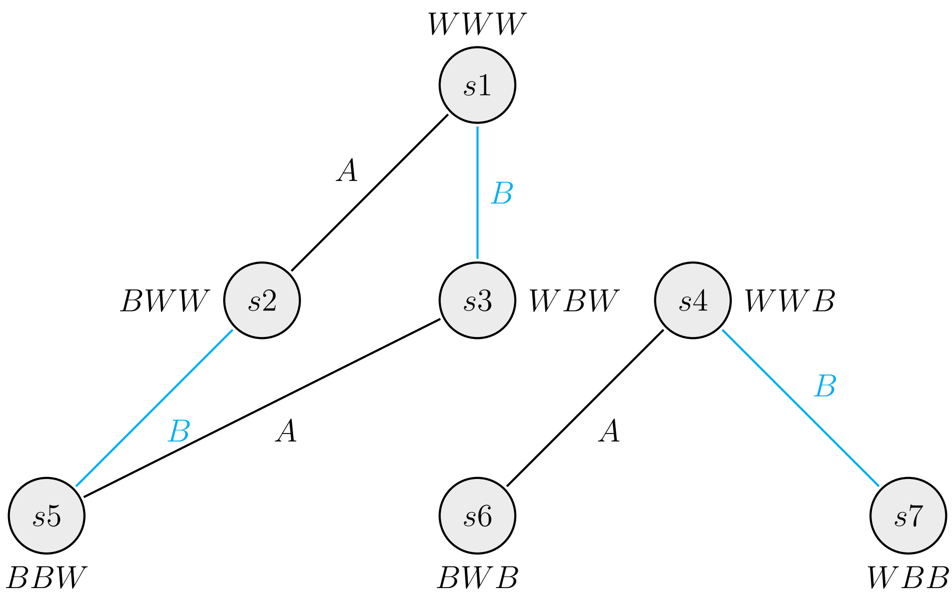

Figure 1: M1

Θα ασχοληθούμε με το ερώτημα β'. Η λίστα με όλους τους πιθανούς κόσμους φαίνεται στον πίνακα 1. Το αντίστοιχο διάγραμμα Kripke φαίνεται στο σχήμα 1. Εφόσον είναι κοινή γνώση πως ο παίκτης C είναι τυφλός, όλοι οι πιθανοί κόσμοι είναι ισοδύναμοι για τον παίκτη C.

Επομένως υπάρχουν ακμές για τον C από κάθε κόσμο σε κάθε άλλο κόσμο. Οι ακμές αυτές δεν αναπαρίστανται για λόγους ευκρίνειας. Παρατηρούμε πως $(K_A \ AisWh)$ αληθεύει μόνο στην κατάσταση $s7$, ενώ $(K_B \ BisWh)$ ισχύει μόνο στην κατάσταση $s6$.

Θα διακρίνουμε τώρα δύο περιπτώσεις. Αν στην πρώτη ερώτηση του βασιλιά είτε ο A, είτε ο B απαντήσουν πως ξέρουν το χρώμα του καπέλου τους, τότε βρισκόμαστε είτε στην κατάσταση $s7$, είτε στην κατάσταση $s6$ αντίστοιχα. Παρατηρούμε τώρα πως ο παίκτης C έχει μαύρο καπέλο και στις δύο αυτές καταστάσεις. Άρα όλοι οι παίκτες ξέρουν το χρώμα τους μετά τη δεύτερη ερώτηση του βασιλιά. Αντιθέτως αν όλοι απαντήσουν αρνητικά στην πρώτη ερώτηση, τότε το διάγραμμα Kripke αλλάζει όπως φαίνεται στο σχήμα 2 (Οι ακμές του παίκτη C παραλείπονται για λόγους ευκρίνειας).

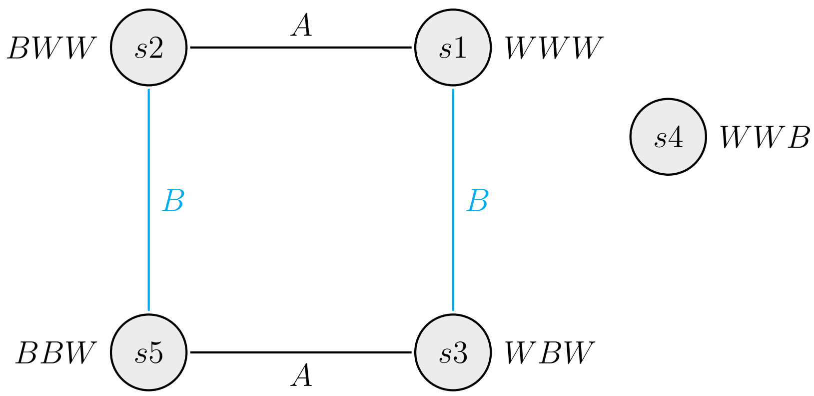

Figure 2: M2

Παρατηρούμε τώρα ότι $(K_A \ AisWh \wedge K_B \ BisWh)$ αληθεύει μόνο στην κατάσταση $s4$. Θα διακρίνουμε όπως πριν δύο περιπτώσεις. Αν στη δεύτερη ερώτηση ο A και ο B απαντήσουν θετικά, τότε ο παίκτης C μαθαίνει πως φοράει μαύρο καπέλο και μπορεί να απαντήσει θετικά στην τρίτη ερώτηση. Αντιθέτως αν όλοι απαντήσουν αρνητικά στη δεύτερη ερώτηση, τότε το διάγραμμα Kripke αλλάζει όπως φαίνεται στο σχήμα 3. Παρατηρούμε όμως πως $(CisWh)$ αληθεύει σε όλες τις πιθανές καταστάσεις του διαγράμματος, κι έτσι ο παίκτης C θα απαντήσει θετικά στην τρίτη ερώτηση. Ταυτόχρονα όμως, επειδή $(CisWh)$ αληθεύει σε όλες τις πιθανές καταστάσεις, ισχύει ότι $(C \ CisWh)$. Σε αυτή την περίπτωση οι παίκτες A και B δε θα μάθουν ποτέ το χρώμα του καπέλου τους, όσες φορές κι αν ρωτήσει ο βασιλιάς. Το διάγραμμα Kripke μένει στάσιμο από εδώ και στο εξής γιατί κανένας παίκτης δεν αποκτά καινούργια πληροφορία.

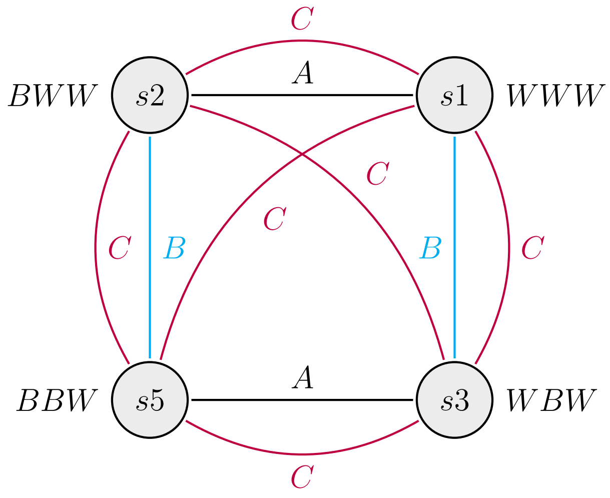

Figure 3: M3

Απάντηση για διαδοχικές αποκρίσεις

Στην περίπτωση που οι παίκτες αποκρίνονται διαδοχικά το αρχικό διάγραμμα Kripke είναι ξανά το σχήμα 1. Παρατηρούμε, όπως και προηγουμένως, πως $(K_A\ AisWh)$ αληθεύει μόνο στην κατάσταση $s7$, ενώ $(K_B\ BisWh)$ ισχύει μόνο στην κατάσταση $s6$. Αν ο παίκτης A απαντήσει θετικά, τότε όλοι οι παίκτες ξέρουν πως βρισκόμαστε στην κατάσταση $s7$, και άρα όλοι μαθαίνουν το χρώμα τους. Αντιθέτως, αν ο παίκτης A απαντήσει αρνητικά, τότε το διάγραμμα Kripke αλλάζει όπως φαίνεται στο σχήμα 4 (Οι ακμές του παίκτη C παραλείπονται για λόγους ευκρίνειας).

Παρατηρούμε τώρα ότι $(K_B\ BisWh)$ αληθεύει στις καταστάσεις $s4$ και $s6$. Θα διακρίνουμε δύο περιπτώσεις. Αν ο B απαντήσει αρνητικά, τότε το διάγραμμα Kripke αλλάζει όπως φαίνεται στο σχήμα 3. Όμως, όπως και πριν, $(K_C\ CisWh)$ ισχύει σε όλες τις καταστάσεις του σχήματος 3. Έτσι λοιπόν, ο παίκτης C μαθαίνει το χρώμα του και απαντά θετικά.

Αντιθέτως, αν ο παίκτης B απαντήσει θετικά, τότε το διάγραμμα Kripke αλλάζει όπως φαίνεται στο σχήμα 5. Παρατηρούμε ότι $(K_C\ \neg CisWh)$ ισχύει σε όλες τις καταστάσεις του σχήματος 5. Έτσι, ο παίκτης C μαθαίνει το χρώμα του και απαντά θετικά. Ο παίκτης A δεν μαθαίνει κάποια καινούργια πληροφορία μετά την απόκριση του C κι έτσι δε θα μάθει ποτέ το χρώμα του, όσες φορές κι αν ρωτήσει ο βασιλιάς.

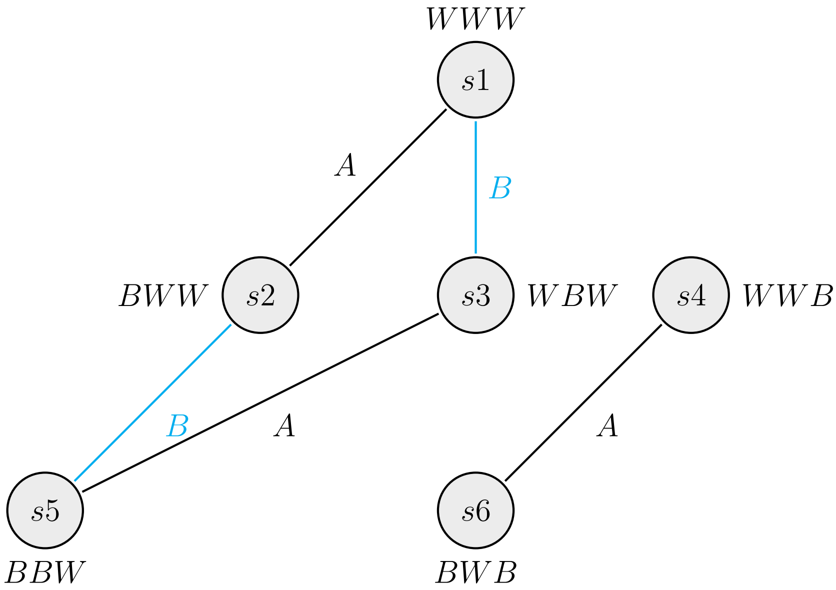

Figure 4: M4

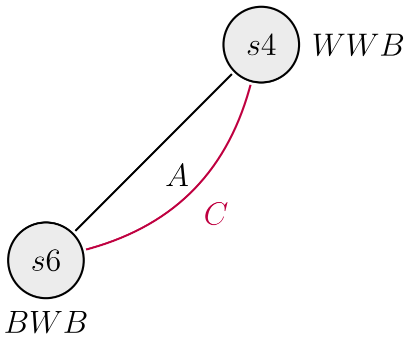

Figure 5: M5

Συμπέρασμα

Συμπερασματικά, στην περίπτωση που είναι κοινή γνώση πως ο παίκτης C είναι τυφλός, τότε το αν αυτός θα μπορέσει να μάθει το χρώμα του εξαρτάται από τον τρόπο που αποκρίνονται οι παίκτες. Πιο συγκεκριμένα, αν οι παίκτες αποκρίνονται διαδοχικά (πρώτα ο A, μετά ο B, μετά ο C), τότε ο παίκτης C θα μπορέσει να απαντήσει θετικά με την πρώτη ερώτηση του βασιλιά. Αν όμως οι παίκτες αποκρίνονται ταυτόχρονα, τότε ο παίκτης C μπορεί να μη μπορέσει να αποκριθεί με την πρώτη ερώτηση του βασιλιά, και να χρειαστεί άλλες δύο ερωτήσεις για να απαντήσει θετικά.