Γενικευμένη Μη Πληρότητα:

Κατανόηση και Παρανόηση των Θεωρημάτων Μη Πληρότητας του Gödel[*]

Α. Αραγεώργης
Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών
Εθνικό Μετσόβιο Πολυτεχνείο
arage@central.ntua.gr

0. Εισαγωγή

Είναι κοινή πεποίθηση ότι τα θεωρήματα μη πληρότητας του Kurt Gödel συγκαταλέγονται ανάμεσα στα πλέον εμβριθή και γοητευτικά διανοητικά επιτεύγματα του 20ου αιώνα. Η πεποίθηση αυτή, σε συνδυασμό με το γεγονός ότι πρόκειται για αποτελέσματα που φέρουν τα διακριτικά της *μαθηματικής βεβαιότητας*, επεφύλαξε στα θεωρήματα μη πληρότητας τον ρόλο των τεκμηρίων για πλήθος ετερόκλητων φιλοσοφικών ισχυρισμών που αφορούν τη θεμελίωση και τη φύση των μαθηματικών, τη σχετική οριοθέτηση λογικής-μαθηματικών-φυσικής, τη «μηχανοποίηση» της ανθρώπινη νόησης, τα όρια της ανθρώπινης γνώσης, την «αποδόμηση» του θεωρητικού λόγου εν γένει ή ακόμη και την χειραφέτηση του ανθρωπίνου είδους και την ύπαρξη του Θεού.

Ο Gödel απέδειξε τα θεωρήματα μη πληρότητας το 1931.[1] Ένα χρόνο πριν είχε αποδείξει το θεώρημα πληρότητας για τον πρωτοβάθμιο κατηγορηματικό λογισμό – σε μια εκδοχή που περιείχαν τα *Principia Mathematica* των Bertrand Russell και Alfred North Whitehead όπως την είχαν επεξεργαστεί οι David Hilbert και Wilhelm Ackermann στο έργο τους *Αρχές της Θεωρητικής Λογικής* [*Grundzüge der theoretischen Logik*, 1928].[2] Ωστόσο, τα θεωρήματα μη πληρότητας είναι αυτά που έχουν ελκύσει περισσότερο το γενικό ενδιαφέρον. Ίσως ακριβώς επειδή είναι *αρνητικά* αποτελέσματα: φαίνεται να περιγράφουν τα *όρια* του τι μπορούμε «να κάνουμε» με τη λογική και τα μαθηματικά.

Φυσικά, μια διεξοδική και λεπτομερής παρουσίαση των *ενδεχόμενων* συνεπειών των θεωρημάτων μη πληρότητας θα ήταν ένας τιτάνιος στόχος.[3] Θα περιοριστώ σε μια σκιαγράφηση μερικών από αυτές αφού πρώτα περιγράψω το περιεχόμενο των θεωρημάτων. Αυτό επιβάλλει μια περιήγηση σε ορισμένες τεχνικές λεπτομέρειες της λογικής και της θεωρίας υπολογισμού. Όμως η εξοικείωση με το ακριβές περιεχόμενο των θεωρημάτων, αλλά και με τον τρόπο απόδειξής τους, είναι απαραίτητη για τη διερεύνηση των συνεπειών τους. Και τούτο όχι χάριν μιας άκρατης προσήλωσης στην καθολική δεοντολογική αρχή «Πρέπει πάντα να γνωρίζεις ακριβώς αυτό για το οποίο μιλάς!» Τα θεωρήματα μη πληρότητας είναι πολύ «ευαίσθητα» στην εκλαΐκευση με την έννοια ότι αόριστες διατυπώσεις οδηγούν σε παραδοξολογίες και σοβαρές παρανοήσεις.

Για να πάρουμε μια γεύση από τις δυσκολίες που απειλούν κάθε επιφανειακή προσέγγιση των θεωρημάτων μη πληρότητας, ας θεωρήσουμε την πρόταση:

(1) «Αυτή η πρόταση δεν μπορεί να αποδειχθεί.»

Είναι αυτή η πρόταση παραδοξολογική; Φαίνεται να είναι. Γιατί αν είναι ψευδής –αν είναι ψευδές ότι δεν μπορεί να αποδειχθεί– τότε *μπορεί* να αποδειχθεί, πράγμα που σημαίνει ότι είναι αληθής. Έτσι αν είναι ψευδής, θα είναι αληθής – άτοπο! Επομένως δεν μπορεί να είναι ψευδής: πρέπει να είναι αληθής. Μόλις αποδείξαμε λοιπόν ότι η πρόταση είναι αληθής. Αλλά αφού είναι αληθής, αυτό που λέει θα ισχύει – δηλαδή, δεν θα μπορεί να αποδειχθεί. Οπότε πώς κατορθώσαμε και την αποδείξαμε;

Η «αύρα» της παραδοξότητας σχετίζεται με *δυο «μυστήρια»* που απασχόλησαν, και εξακολουθούν να απασχολούν, τους φιλοσόφους από την εποχή των προσωκρατικών. Πρόκειται για την *άρνηση* και την *αυτο-αναφορά*. Το παράδοξο επιτείνεται αν προσθέσει κανείς εσφαλμένα –όπως, δυστυχώς, συμβαίνει συχνά– ότι η απόδειξη του πρώτου θεωρήματος μη πληρότητας από τον Gödel βασίζεται στην παραπάνω πρόταση. *Πώς είναι δυνατόν ένα μη τετριμμένο θεώρημα να βασίζεται σε μια «αντίφαση»*; Αλλά το παράδοξο διαλύεται –μάλλον, *αρχίζει* να διαλύεται– μόλις κατανοήσει κανείς (α) ότι στο παραπάνω επιχείρημα η έννοια της απόδειξης δεν είναι καλώς ορισμένη και ότι πρέπει κανείς να μιλά μάλλον για *αποδειξιμότητα στο πλαίσιο ενός τυπικού συστήματος* και (β) ότι στα τυπικά συστήματα της κλασικής λογικής δεν ισχύει ότι οτιδήποτε μπορεί να αποδειχθεί από κάποιο μη κενό σύνολο προκειμένων είναι και αληθές με κάποια «απόλυτη» σημασία.[4]

Τα θεωρήματα μη πληρότητας του Gödel είναι αυστηρά αποτελέσματα για την *πρωτοβάθμια λογική* και για αξιωματικές θεωρίες που μπορούν να διατυπωθούν σε αυτή, όπως η αριθμητική Peano και η θεωρία συνόλων Zermelo-Fraenkel. Και ως τέτοια πρέπει να αντιμετωπίζονται. Αλλιώς θα καταλήγουμε σε δημοφιλείς αλλά παραπλανητικές αοριστίες όπως ο ισχυρισμός: «Όπως απέδειξε ο Gödel, υπάρχουν αλήθειες που δεν αποδεικνύονται!» Και τέτοιες αοριστίες είναι εξίσου εσφαλμένες όσο ο ισχυρισμός: «Όπως απέδειξε ο Einstein, τα πάντα είναι σχετικά!» Η απρόσεκτη εκλαΐκευση και η επιφανειακή πραγμάτευση επεφύλαξαν, στο πλαίσιο γενικών συζητήσεων, ανάλογη κακομεταχείριση στα θεωρήματα μη πληρότητας του Kurt Gödel και στη θεωρία της σχετικότητας του Albert Einstein – στα πνευματικά προϊόντα δυο μεγάλων στοχαστών του 20ου αιώνα που έτυχε να συνδεθούν φιλικά κατά τη διάρκεια της παραμονής τους στο Ινστιτούτο Ανωτάτων Σπουδών του Princeton.

1. Τα θεωρήματα μη πληρότητας[5]

Ας φανταστούμε μια γλώσσα L που είναι εφοδιασμένη με ένα κατάλληλο σύστημα λογικών αξιωμάτων και αποδεικτικών κανόνων. Μια *θεωρία* T στη γλώσσα L είναι ένα σύνολο προτάσεων της L τέτοιο ώστε για κάθε πρόταση φ της L, αν η φ *αποδεικνύεται από την* (ή *είναι αποδείξιμη στην*) T (συμβολικά, $T \vdash \varphi$), τότε η φ ανήκει στην T (συμβολικά, $\varphi \in T$). Η θεωρία T λέγεται *συνεπής* αν και μόνον αν δεν υπάρχει πρόταση φ της L τέτοια ώστε $T \vdash \varphi$ και $T \vdash \neg\varphi$, όπου εδώ και στο εξής με '$\neg\varphi$' συμβολίζουμε την άρνηση μιας πρότασης φ της L - αλλιώς, η θεωρία T θα αποδείκνυε και, συνεπώς, θα περιείχε τη λογική αντίφαση $\varphi \wedge \neg\varphi$, όπου με '$\wedge$' συμβολίζουμε τον λογικό σύνδεσμο της σύζευξης ('και') στην L. Η θεωρία T λέγεται *πλήρης* ακριβώς στην περίπτωση για κάθε πρόταση φ της L είτε $T \vdash \varphi$ είτε $T \vdash \neg\varphi$. Έπεται ότι η θεωρία T θα λέγεται *μη πλήρης* αν και μόνον αν υπάρχει μια πρόταση φ της γλώσσας της θεωρίας που ούτε η ίδια ούτε η άρνησή της είναι αποδείξιμη στην θεωρία (συμβολικά, $T \nvdash \varphi$ και $T \nvdash \neg\varphi$, αντίστοιχα) – με αυτή την έννοια, η θεωρία T δεν αποφαίνεται ούτε για την «αλήθεια» ούτε για το «ψεύδος» της φ και για αυτό η πρόταση φ χαρακτηρίζεται ως *μη αποφασίσιμη*, *μη αποκρίσιμη* ή *μη απαντήσιμη* στην T.

Μια θεωρία T στη γλώσσα L λέγεται *αξιωματικοποιήσιμη* αν και μόνο αν υπάρχει ένα σύνολο A προτάσεων της L («αξιώματα» της T) τέτοιο ώστε (α) για κάθε πρόταση φ της L, $T \vdash \varphi$ ακριβώς στην περίπτωση που $A \vdash \varphi$ και (β) υπάρχει *αλγόριθμος* που αποφασίζει για οποιαδήποτε πρόταση της L εάν ανήκει στο A ή όχι.

Αλλά *τι ακριβώς είναι ένας αλγόριθμος*; Η έννοια του αλγορίθμου φαίνεται να είναι τόσο αρχαία όσο και τα μαθηματικά. Οι αρχαίοι Έλληνες γεωμέτρες επινοούσαν σταδιακές διαδικασίες κατασκευής γεωμετρικών αντικειμένων με κανόνα και διαβήτη. Ακόμη και σήμερα ένα μεγάλο μέρος της βασικής μαθηματικής εκπαίδευσης συνίσταται στη διδαχή αλγορίθμων για την επίλυση προβλημάτων – π.χ., πρόσθεση ή πολλαπλασιασμός αριθμών με πολλά ψηφία, λύση αλγεβρικών εξισώσεων 1ου βαθμού, κ.λπ. Παρ' όλα αυτά η έννοια του αλγορίθμου δεν έχει ακόμη κατανοηθεί πλήρως: *δεν υπάρχει ένας καθολικά αποδεκτός ορισμός του όρου 'αλγόριθμος'*.[6] Μπορεί, βέβαια, να καταφύγει σε γενικόλογες περιγραφές όπως η ακόλουθη:[7]

(2) Ένας αλγόριθμος είναι μια αλάθητη, μηχανική, βήμα προς βήμα διαδικασία η οποία, τουλάχιστον ιδεατά, μπορεί να επιλύσει ένας είδος προβλήματος για οποιαδήποτε δεδομένα από ένα (άπειρο) σύνολο δυνατών δεδομένων.

Με τη λέξη '*αλάθητη*' εννοούμε ότι είναι εγγυημένη η επίλυση του προβλήματος με πεπερασμένο αριθμό βημάτων (εφόσον εκτελεστούν σωστά). Με τη λέξη '*μηχανική*' εννοούμε ότι ένας ιδεατός υπολογιστής μπορεί να προγραμματιστεί να την εκτελέσει. Με την έκφραση '*βήμα προς βήμα*' εννοούμε ότι η διαδικασία ορίζει ένα βήμα κάθε φορά, το ένα μετά το άλλο, και ότι έπειτα από κάθε βήμα το επόμενο είναι όχι μόνο καθορισμένο αλλά και φανερό - δηλαδή, όχι μόνο δεν υπάρχουν αβεβαιότητες αλλά και δεν απαιτείται έμπνευση για να ανακαλυφθεί ποιο πρέπει να είναι το επόμενο βήμα. Τέλος, χρησιμοποιούμε την έκφραση '*ιδεατά*' για να δηλώσουμε ότι αγνοούμε τυχόν ύπαρξη σταθερών περιορισμών χρόνου, μεγέθους μνήμης, κ.λπ.

Όμως μόνο με γενικόλογες περιγραφές σαν τη (2) παραπάνω δεν αποδεικνύονται *μαθηματικά θεωρήματα*. Οπότε είναι αναγκαίο να δούμε με μεγαλύτερη λεπτομέρεια ποια έννοια αλγορίθμου υπεισέρχεται στα θεωρήματα μη πληρότητας και με ποιο τρόπο. Ας θεωρήσουμε πάλι την τυπική γλώσσα L και ας υποθέσουμε ότι θέλουμε να διαπιστώσουμε εάν μια αυθαίρετα επιλεγμένη πρόταση φ της L ανήκει σε ένα υποσύνολο A του συνόλου των προτάσεων της L ή όχι. Πώς μπορεί το πρόβλημα αυτό να αναχθεί σε πρόβλημα υπολογισμού; Το κλειδί είναι η *αριθμητικοποίηση της σύνταξης* της L: σε κάθε έκφραση σ της L αποδίδουμε ένα φυσικό αριθμό $g(\sigma)$, τον οποίο ονομάζουμε *αριθμό Gödel* της σ, έτσι ώστε (α) σε διαφορετικές εκφράσεις να αποδίδονται διαφορετικοί αριθμοί Gödel (β) να μπορεί να υπολογιστεί αποτελεσματικά ποιος είναι ο αριθμός Gödel οποιασδήποτε έκφρασης, και (γ) δεδομένου οποιουδήποτε αριθμού να μπορεί να αποφασιστεί αποτελεσματικά εάν ο αριθμός αυτός είναι ο αριθμός Gödel κάποιας έκφρασης και, αν είναι, να βρεθεί αποτελεσματικά ποια είναι η έκφραση αυτή. Με αυτό τον τρόπο, το αρχικό ερώτημα «Ανήκει η πρόταση φ στο σύνολο A;» μετασχηματίζεται στο ερώτημα «Ανήκει ο αριθμός Gödel $g(\varphi)$ της φ στο σύνολο $g[A]=\{g(\sigma):\sigma\in A\}$ των αριθμών Gödel των στοιχείων του συνόλου A;» Και το δεύτερο ερώτημα μετασχηματίζεται, ένα βήμα παραπέρα, στο ερώτημα «Απεικονίζει η χαρακτηριστική συνάρτηση $\chi_{g[A]}$ του συνόλου $g[A]$ το όρισμα $g(\varphi)$ στην τιμή 1;»[8] Με άλλα λόγια, για να αποφασίσουμε εάν η φ ανήκει στο A ή όχι αρκεί πλέον να υπολογίσουμε την τιμή $\chi_{g[A]}(g(\varphi))$: αν η τιμή ισούται με 0, τότε $\varphi\notin A$ ενώ αν η τιμή ισούται με 1, τότε $\varphi\in A$. Το ζήτημα τώρα είναι εάν υπάρχει αλγόριθμος για τον υπολογισμό της $\chi_{g[A]}$, εάν η συνάρτηση $\chi_{g[A]}$ είναι *υπολογιστή*. Αλλά πώς ορίζεται η κλάση των υπολογιστών συναρτήσεων; Μαθηματικοί και ειδικοί της λογικής έχουν επεξεργαστεί διάφορα πρότυπα υπολογιστών συναρτήσεων (συναρτήσεις υπολογιστές από μηχανές Turing, αναδρομικές συναρτήσεις, κ.ά.) και έχουν αποδείξει ότι πολλά από αυτά τα διαισθητικά εύλογα πρότυπα είναι εκτασιακά ισοδύναμα (δηλαδή, ότι μια συνάρτηση μπορεί να υπολογιστεί από κάποια μηχανή Turing αν και μόνο αν είναι αναδρομική, κ.λπ.) Αυτές τα αποτελέσματα έχουν εκληφθεί ως τεκμήρια υπέρ της υπόθεσης που έχει γίνει γνωστή ως *θέση του Church*: οι υπολογιστές αριθμητικές συναρτήσεις είναι ακριβώς εκείνες που μπορούν να υπολογιστούν από μηχανές Turing. Από αυτή τη σκοπιά, οι μηχανές Turing προσφέρουν μια αυστηρή μαθηματική επεξήγηση της έννοιας του αλγορίθμου: υπάρχει αλγόριθμος για τον υπολογισμό των τιμών μιας συνάρτησης για τα διάφορα ορίσματα αν και μόνο αν υπάρχει μια μηχανή Turing που υπολογίζει τις τιμές αυτές. Ας αποδώσουμε, στο εξής, τον χαρακτηρισμό ‘*αλγοριθμικές συναρτήσεις*’ σε όλες τις υπολογιστές κατά Turing συναρτήσεις (και μόνο σε αυτές) και ας αποδώσουμε τον χαρακτηρισμό ‘*αλγοριθμικό σύνολο*’ σε όλα τα υποσύνολα του συνόλου των φυσικών αριθμών των οποίων οι χαρακτηριστικές συναρτήσεις είναι αλγοριθμικές (και μόνο σε αυτά).

Με τη βοήθεια της θέσης του Church, λοιπόν, έχουμε έναν ακριβή ορισμό της αλγοριθμικότητας που μπορεί να αξιοποιηθεί στη συναγωγή μαθηματικών θεωρημάτων. Ωστόσο δεν πρέπει να λησμονούμε ότι η ίδια η θέση του Church *δεν* *είναι μαθηματικό αποτέλεσμα*· είναι μάλλον μια *εισήγηση για ορισμό*. Η εισήγηση έχει ως εξής: λαμβάνοντας υπόψη (α) ότι διαφορετικές εννοιολογικές προσεγγίσεις της έννοιας της υπολογιστής συνάρτησης συγκλίνουν σε *μια* κλάση συναρτήσεων, (β) ότι κάθε μέλος αυτής της κλάσης συναρτήσεων μπορεί να υπολογιστεί από απλές μηχανές και (γ) ότι κάθε συνάρτηση που ο καθένας μας διαισθητικά αντιλαμβάνεται ως υπολογιστή είναι υπολογιστή κατά Turing είναι εύλογο να εκλάβουμε ως υπολογιστές συναρτήσεις ακριβώς εκείνες που είναι υπολογιστές κατά Turing. Επιπλέον, επειδή μπορούμε να αναγάγουμε συναρτήσεις που ορίζονται πάνω σε πεπερασμένες ακολουθίες αντικειμένων οποιουδήποτε είδους σε συναρτήσεις πάνω στους φυσικούς αριθμούς, η θέση του Church έχει ευρύτερες συνέπειες. Δεν αφορά μόνον την υπολογιστότητα συναρτήσεων πάνω στους αριθμούς, αλλά και την υπολογιστότητα συναρτήσεων πάνω σε οποιαδήποτε αντικείμενα μπορούν να απαριθμηθούν με μια 1-1 υπολογιστή απεικόνιση στους φυσικούς αριθμούς. Γενικότερα, αλλά ασαφέστερα, η θέση του Church λέει ότι οτιδήποτε μπορεί να υπολογιστεί μπορεί να αναχθεί στον υπολογισμό μιας συνάρτησης την οποία υπολογίζει κάποια μηχανή Turing (ισοδύναμα, στον υπολογισμό μιας αναδρομικής συνάρτησης, κ.λπ.).

Αλλά ας επιστρέψουμε τώρα στα θεωρήματα μη πληρότητας του Gödel. Στις συνηθισμένες διατυπώσεις τους, τα θεωρήματα αυτά φέρονται να αφορούν τυπικές μαθηματικές θεωρίες που «περιλαμβάνουν» την αριθμητική – ακριβέστερα, πρωτοβάθμιες θεωρίες που επεκτείνουν την αριθμητική Peano. Πολύ γενικά, η *αριθμητική Peano* (στο εξής, PA για συντομία) είναι μια αξιωματική θεωρία σε μια πρωτοβάθμια γλώσσα της οποίας το πρότυπο μοντέλο είναι η «στοιχειώδης» αριθμητική των φυσικών αριθμών, δηλαδή η δομή $Mod_{\mathrm{PA}} = \langle \mathbb{N}, s, +, \cdot, 0 \rangle$, όπου $\mathbb{N} = \{0,1,2,\ldots\}$ το σύνολο των φυσικών αριθμών, s η συνάρτηση που απεικονίζει κάθε φυσικό αριθμό στον επόμενό του, $+$ η πράξη της πρόσθεσης μεταξύ φυσικών αριθμών, $\cdot$ η πράξη του πολλαπλασιασμού μεταξύ φυσικών αριθμών και 0 ο «ελάχιστος» φυσικός αριθμός. Θα λέμε ότι *μια θεωρία* T *σε μια γλώσσα* L *επεκτείνει την* PA αν και μόνο αν η L περιέχει όλα τα σύμβολα της γλώσσας της PA και κάθε αξίωμα της PA είναι αποδείξιμο στην T.

Το πρώτο θεώρημα μη πληρότητας του Gödel (στο εξής, 1οΘΜΠ) μπορεί πλέον να διατυπωθεί ως εξής:

(3) 1οΘΜΠ. Κάθε συνεπής αξιωματικοποιήσιμη θεωρία που επεκτείνει την PA είναι μη πλήρης.

Ας σημειωθεί ότι το θεώρημα γενικεύεται για όλα τα τυπικά συστήματα στα οποία είναι αναπαραστάσιμες όλες οι αναδρομικές συναρτήσεις φυσικών αριθμών. Ένα *ασθενέστερο* από την PA τέτοιο σύστημα είναι η *αριθμητική Robinson* που, ουσιαστικά, είναι η PA χωρίς το αξιωματικό σχήμα της μαθηματικής επαγωγής αλλά με προσθήκη σε θέση αξιώματος της πρότασης ότι κάθε αριθμός διάφορος του μηδενός έχει επόμενο.[9] Φυσικά, το θεώρημα ισχύει για όλα τα τυπικά συστήματα που είναι *ισχυρότερα* από την PA, όπως η αξιωματική θεωρία συνόλων Zermelo-Fraenkel με αξίωμα επιλογής (ZFC) που διεκδικεί τον τίτλο του θεμελίου όλων των κλασικών μαθηματικών.

Η απόδειξη του 1ουΘΜΠ περιλαμβάνει δυο στάδια. Ας υποθέσουμε ότι έχουμε μια αξιωματικοποιήσιμη θεωρία T που επεκτείνει την PA. Και ας υποθέσουμε ότι έχουμε αριθμητικοποιήσει τη σύνταξη της γλώσσας L της T. *Πρώτα*, αποδεικνύουμε ότι υπάρχει κατηγόρημα P στην L τέτοιο ώστε για κάθε πρόταση φ της L,

(4) $T \vdash \varphi$ αν και μόνο αν $T \vdash P(g(\varphi))$.

Δηλαδή, μια πρόταση φ αποδεικνύεται στη θεωρία T αν και μόνο αν αποδεικνύεται στη θεωρία T η πρόταση που προκύπτει με απόδοση του κατηγορήματος P στον αριθμό Gödel της φ. Με αυτή την έννοια, το κατηγόρημα P *κωδικοποιεί την ιδιότητα της T-αποδειξιμότητας*. Στο δεύτερο στάδιο της απόδειξης, κατασκευάζουμε μια πρόταση G_T («πρόταση Gödel για την T») που *στο πλαίσιο της T* ισοδυναμεί με την άρνηση της αποδειξιμότητάς της:

(5) $T \vdash G_T \leftrightarrow \neg P(g(G_T))$.

(Εδώ το '$\leftrightarrow$' συμβολίζει το λογικό σύνδεσμο της υλικής ισοδυναμίας στη γλώσσα L.) Αυτό σημαίνει ότι η G_T είναι T-αποδείξιμη αν και μόνο αν η $\neg P(g(G_T))$ είναι T-αποδείξιμη και ότι η $\neg G_T$ είναι T-αποδείξιμη αν και μόνο αν η $P(g(G_T))$ είναι

T-αποδείξιμη. Ας υποθέσουμε τώρα ότι πρόταση G_T είναι αποφασίσιμη στην T - δηλαδή, ότι η T αποδεικνύει είτε τη G_T είτε τη $\neg G_T$. Έστω ότι η G_T αποδεικνύεται στην T. Τότε, σύμφωνα με το (4), η $P(g(G_T))$ θα είναι επίσης T-αποδείξιμη και, σύμφωνα με το (5), η $\neg P(g(G_T))$ θα είναι επίσης T-αποδείξιμη. Επομένως η θεωρία T θα είναι ασυνεπής αφού θα αποδεικνύει και την $P(g(G_T))$ και την άρνησή της. Έστω τώρα ότι η $\neg G_T$ αποδεικνύεται στην T. Τότε, σύμφωνα με το (5), η $P(g(G_T))$ θα είναι επίσης T-αποδείξιμη οπότε, σύμφωνα με το (4), η G_T θα είναι επίσης T-αποδείξιμη. Επομένως η θεωρία T θα είναι πάλι ασυνεπής αφού θα αποδεικνύει και την G_T και την άρνησή της. Άρα αν η πρόταση G_T είναι αποφασίσιμη στην T, η θεωρία T θα είναι ασυνεπής. Έπεται ότι αν η θεωρία T είναι συνεπής, η πρόταση G_T είναι μη αποφασίσιμη στην T και η T είναι μη πλήρης.

Αυτό ολοκληρώνει τη σκιαγράφηση των κομβικών σημείων μιας απόδειξης του 1ουΘΜΠ. Όπως υποψιάζεται κάθε αναγνώστης, οι λεπτομερείς αποδείξεις των δυο σταδίων που απλώς αναφέρθηκαν παραπάνω απαιτούν σημαντική προσπάθεια ενώ η αρχική επινόησή τους από τον Kurt Gödel απαιτούσε ιδιαίτερη ευφυΐα και έμπνευση. Πρέπει επίσης να σημειώσω ότι ο τρόπος απόδειξης που σκιαγράφησα προϋποθέτει επιπλέον ότι η θεωρία T είναι *ω-συνεπής*.[10] Η ω-συνέπεια της T είναι μια τεχνική υπόθεση την οποία είχε χρησιμοποιήσει ρητά ο Gödel ([1931] 1967, σ. 607) στην αρχική απόδειξη. Την αποσιώπησα για λόγους απλούστευσης του κυρίως κειμένου δεδομένου ότι την κατέστησε περιττή μια *τροποποιημένη* απόδειξη του 1ουΘΜΠ στη διατύπωση (3) που κατασκεύασε ο J. Barkley Rosser το 1936. Ωστόσο είναι σημαντικό και συνάμα διδακτικό το γεγονός ότι, όπως θα διαπιστώσουμε παρακάτω, ακόμη και τέτοιες «τόσο τεχνικές λεπτομέρειες» της απόδειξης ενδέχεται να παίξουν κρίσιμο ρόλο στη διερεύνηση των συνεπειών του 1ουΘΜΠ.

Το δεύτερο θεώρημα μη πληρότητας (στο εξής, 2οΘΜΠ) προκύπτει ως συνέπεια του 1ουΘΜΠ. Αποδεικνύεται ότι για κάθε συνεπή αξιωματικοποιήσιμη πρωτοβάθμια θεωρία T που επεκτείνει την PA, η πρόταση $Cons(T) = \neg P(g(\perp))$, όπου P το κατηγόρημα που κωδικοποιεί την T-αποδειξιμότητα και $\perp$ μια αντίφαση στη γλώσσα της T (π.χ., η $\neg(1=1)$) είναι ισοδύναμη *στο πλαίσιο της* T με τη πρόταση Gödel G_T για την T:

(6) $T \vdash G_T \leftrightarrow Cons(T)$, όπου $Cons(T) = \neg P(g(\bot))$.

Οπότε αφού, από το 1οΘΜΠ, η G_T είναι μη αποφασίσιμη στη θεωρία T, μη αποφασίσιμη στη θεωρία T θα είναι και η πρόταση $Cons(T)$. Συνοψίζοντας, έχουμε:

(7) 2οΘΜΠ. Αν T είναι μια συνεπής αξιωματικοποιήσιμη θεωρία που επεκτείνει την PA, P είναι το κατηγόρημα που κωδικοποιεί την T-αποδειξιμότητα και $Cons(T) = \neg P(g(\bot))$ όπου $\bot$ μια αντίφαση στη γλώσσα της T, τότε $T \nvdash Cons(T)$ και $T \nvdash \neg Cons(T)$.

Εφόσον η $Cons(T)$ εκφράζει την αδυναμία της T να αποδείξει μια αντίφαση και, με αυτή τον τρόπο, τη *συνέπεια* της T, το 2οΘΜΠ μη πληρότητας συνεπάγεται ότι καμιά συνεπής αξιωματικοποιήσιμη θεωρία που επεκτείνει την PA δεν μπορεί να αποδείξει την ίδια της τη συνέπεια. Γενικεύοντας, *κανένα συνεπές αξιωματικό σύστημα που είναι αρκετά ισχυρό ώστε να μπορεί να αναπαραστήσει κάθε αναδρομική σχέση μεταξύ φυσικών αριθμών δεν μπορεί να αποδείξει την ίδια του τη συνέπεια.* Παρατηρήστε, ωστόσο, ότι έχω δώσει έμφαση στην *υποθετική* φύση της πρότασης με την οποία διατυπώνω την καθιερωμένη ερμηνεία της $Cons(T)$. Όπως θα δούμε, αυτό έχει σημασία στην αποτίμηση των συνεπειών του 2ουΘΜΠ στη φιλοσοφία των μαθηματικών. Αλλά, πρώτα, ας περάσουμε σε μερικές επιγραμματικές διευκρινίσεις για τις συνέπειες των θεωρημάτων μη πληρότητας.

2. Τι δεν λένε τα θεωρήματα μη πληρότητας

Στην εκλαϊκευτική, κυρίως, βιβλιογραφία απαντούν διάφοροι εσφαλμένοι ισχυρισμοί σχετικά με το «τι λένε» για τα μαθηματικά τα θεωρήματα μη πληρότητας. Στη σύντομη αυτή ενότητα, σταχυολογώ μερικούς από αυτούς και δικαιολογώ επιγραμματικά τη θέση ότι είναι εσφαλμένοι.

Ισχυρισμός: «Κάθε ενδιαφέρον συνεπές αξιωματικό σύστημα είναι μη πλήρες». *Λάθος!* Η ευκλείδεια γεωμετρία (π.χ., στην κατά Hilbert [1899] 1992 αξιωματική θεμελίωση) είναι και συνεπής και πλήρης και, με τεκμήριο δυο χιλιετηρίδες περίπου επιτυχιών, μάλλον ενδιαφέρουσα.

Ισχυρισμός: «Κάθε συνεπές αξιωματικό σύστημα που περιέχει τους φυσικούς αριθμούς είναι μη πλήρες». *Λάθος!* Η αριθμητική Presburger που αποδεικνύει κάθε πρωτοβάθμια πρόταση που αναφέρεται μόνο στην πρόσθεση φυσικών αριθμών είναι και συνεπής και πλήρης.[11]

Ισχυρισμός: «Δεν μπορεί να αποδειχθεί η συνέπεια της αριθμητικής». *Λάθος!* Το 1936 ο Gentzen δημοσίευσε μια απόδειξη της συνέπειας της πρωτοβάθμιας αριθμητικής Peano, της PA, χρησιμοποιώντας, όμως, ισχυρότερα εργαλεία που δεν μπορούσαν να τυποποιηθούν στο πλαίσιο της ίδιας της PA.

Ισχυρισμός: «Κανείς δεν μπορεί να γνωρίζει ότι οι μαθηματικές θεωρίες είναι συνεπείς». *Λάθος!* Έχουμε ήδη αποδείξει τη συνέπεια πολλών μαθηματικών θεωριών. Εκείνο που συνεπάγεται το 2οΘΜΠ είναι ότι κάθε μέλος μιας ευρείας κλάσης πρωτοβάθμιων θεωριών αδυνατεί να αποδείξει την πρόταση που εκφράζει *συντακτικά* τη *δική του* συνέπεια.[12] Αλλά θα το θέλαμε αυτό; Ας πούμε ότι μια θεωρία αποδείκνυε την πρόταση που εκφράζει συντακτικά τη δική της συνέπεια. Θα είχαμε τότε λόγους να πιστέψουμε ότι η θεωρία είναι συνεπής; *Ασφαλώς όχι!* Αν η θεωρία είναι ασυνεπής, τότε αποδεικνύει οποιαδήποτε πρόταση[13] – άρα, και την πρόταση που εκφράζει συντακτικά τη δική της συνέπεια. *Το να πιστεύουμε ότι μια θεωρία είναι συνεπής επειδή αποδεικνύει την ίδια της τη συνέπεια είναι τόσο ανόητο όσο το να πιστεύουμε ότι ένας άνθρωπος είναι ειλικρινής επειδή ισχυρίζεται ότι δεν ψεύδεται ποτέ.*[14]

3. Τετριμμένη μη πληρότητα;

Μήπως, όμως, η μεταμαθηματική ιδιότητα της μη πληρότητας που διαπιστώνουν τα θεωρήματα του Gödel είναι μια απολύτως τετριμμένη ιδιότητα την οποία μοιράζονται πολλά λιγότερο ή περισσότερο τυπικά συστήματα; Σκεφθείτε το εξής. Μια θεωρία είναι πλήρης αν, για κάθε πρόταση στη γλώσσα της θεωρίας, η θεωρία αποδεικνύει ή την ίδια την πρόταση ή την άρνησή της. Και ας θεωρήσουμε ένα μυθιστόρημα ή ένα διήγημα ως μια «θεωρία», ως μια συλλογή προτάσεων, για κάποια άτομα, για τους χαρακτήρες που «εμφανίζονται» σε αυτό. Είναι τα μυθιστορήματα ή τα διηγήματα πλήρη με την έννοια ότι μας επιτρέπουν να αποφανθούμε για την αλήθεια ή το ψεύδος κάθε δυνατής πρότασης που αφορά τους χαρακτήρες στους οποίους αναφέρονται; Γενικά, όχι. Δεν μπορεί, λόγου χάριν, κανείς να μάθει τι ύψος είχε ή τι απέγινε η αρραβωνιαστικιά του Georg Bendemann διαβάζοντας το *The Judgment* του Kafka. Και το ίδιο ισχύει για τα περισσότερα μυθιστορήματα ή διηγήματα που μπορούμε να φανταστούμε.[15]

Ή σκεφθείτε μια φυσική θεωρία, π.χ., τη θεωρία που απαρτίζεται από τους τρεις νόμους του Kepler για την κίνηση των πλανητών. Αποφαίνεται αυτή η θεωρία για την αλήθεια ή το ψεύδος κάθε δυνατής υπόθεσης που αφορά τους πλανήτες του ηλιακού μας συστήματος; Όχι. Δεν προσδιορίζει το πλήθος των πλανητών, την περίοδο ή τη μέση απόσταση από τον ήλιο ενός πλανήτη, κ.λπ. Και το «φαινόμενο» είναι γενικό: ακόμη και εκείνες οι φυσικές θεωρίες που έχουν κατά καιρούς εκληφθεί ως θεμελιώδεις υπολείπονται της πληρότητας με την έννοια ότι αφήνουν αναπάντητα πολλά ερωτήματα – για παράδειγμα, ερωτήματα που αφορούν τις ακριβείς τιμές φυσικών σταθερών.

Ίσως, όμως, αυτές οι παρατηρήσεις να αφήνουν αδιάφορους όσους αναγνώστες πιστεύουν ότι η μη πληρότητα, ως μεταμαθηματική έννοια, πρέπει να εφαρμόζονται μόνο σε περιπτώσεις μαθηματικών θεωριών. Ας επιστρέψουμε, λοιπόν, στα μαθηματικά. Αν από την ευκλείδεια γεωμετρία εξαιρέσουμε το αίτημα των παραλλήλων του Ευκλείδη,[16] τότε θα έχουμε μια μη πλήρη θεωρία. Για παράδειγμα, δεν θα μπορούμε να αποφανθούμε εάν το άθροισμα των εσωτερικών γωνιών ενός τριγώνου είναι μικρότερο από 180^{o} ή όχι. Στη μη ευκλείδεια γεωμετρία του Lobachevski, σύμφωνα με την οποία από σημείο εκτός ευθείας άγονται άπειρες παράλληλες προς την ευθεία αυτή, το άθροισμα είναι μικρότερο των 180^{o}. Στην ευκλείδεια γεωμετρία, σύμφωνα με την οποία από σημείο εκτός ευθείας άγεται ακριβώς μια παράλληλος προς την ευθεία αυτή, το άθροισμα ισούται με 180^{o}. Ενώ στη μη ευκλείδεια γεωμετρία του Riemann, σύμφωνα με την οποία από σημείο εκτός ευθείας δεν άγεται καμία παράλληλος προς την ευθεία αυτή, το άθροισμα είναι μεγαλύτερο από 180^{o}.

Ή, για να χρησιμοποιήσουμε μια απλή αλγεβρική θεωρία, σκεφτείτε τη θεωρία ομάδων. Τα μοντέλα της θεωρίας είναι δομές που απαρτίζονται από ένα μη κενό σύνολο μεταξύ των στοιχείων του οποίου ορίζεται μια διμελής πράξη η οποία είναι προσεταιριστική, έχει ουδέτερο στοιχείο και προβλέπει αντίστροφο στοιχείο για κάθε στοιχείο του συνόλου. Όπως είναι γνωστό, υπάρχουν ομάδες στις οποίες η πράξη είναι αντιμεταθετική (π.χ., το σύνολο $\mathbb{R}-\{0\}$ των μη μηδενικών πραγματικών αριθμών με πράξη το συνήθη πολλαπλασιασμό και ουδέτερο στοιχείο τη μονάδα) καθώς και ομάδες στις οποίες η πράξη δεν είναι αντιμεταθετική (π.χ., το σύνολο $GL(2,\mathbb{R})$ των 2×2 πινάκων πραγματικών αριθμών που έχουν μη μηδενική ορίζουσα με πράξη τον πολλαπλασιασμό πινάκων και ουδέτερο στοιχείο τον μοναδιαίο πίνακα). Έπεται ότι η πρόταση που εκφράζει την αντιμεταθετικότητα της πράξης είναι μη αποφασίσιμη στη θεωρία ομάδων.

Τα παραπάνω παραδείγματα φαίνεται να αναδεικνύουν μια μάλλον τετριμμένη όψη της μη πληρότητας θεωριών. Μη αποφασίσιμες προτάσεις αναδύονται σε όλες τις θεωρίες που υπολείπονται μιας πλήρους περιγραφής του αντικειμένου τους. Γιατί, λοιπόν, θεωρείται τόσο σημαντική η συμβολή των θεωρημάτων του Gödel; Μήπως τα θεωρήματα αυτά καταδεικνύουν μόνον το γεγονός ότι η αξιωματικοποίηση της αριθμητικής κατά Peano είναι ελλιπής με την έννοια ότι απαιτεί την προσθήκη κάποιου επιπλέον αξιώματος για την πλήρη περιγραφή του σύμπαντος των φυσικών αριθμών;

Όχι, ακριβώς. Ας προσθέσουμε ως αξιώματα της αριθμητικής όλες τις αλήθειες για τους φυσικούς αριθμούς – πιο συγκεκριμένα, όλες τις προτάσεις που αληθεύουν στο πρότυπο μοντέλο Mod_{PA}. Η θεωρία που θα πάρουμε θα είναι πλήρης αλλά όχι αξιωματικοποιήσιμη. Το σύνολο των αξιωμάτων της δεν θα είναι αλγοριθμικό. Και σε μια τέτοια θεωρία δεν θα μπορούμε να ελέγξουμε μηχανικά εάν μια πεπερασμένη ακολουθία τύπων είναι απόδειξη ή όχι. Από αυτή την άποψη, τα θεωρήματα μη πληρότητας του Gödel έχουν να κάνουν με την *ανυπαρξία αλγορίθμων*. Και είναι σημαντικό να έχουμε αυτή την άποψη κατά νου όταν διερευνούμε τις συνέπειες των θεωρημάτων αυτών σε κάποια περιοχή της φιλοσοφίας. Στη σύντομη επισκόπηση μερικών τέτοιων συνεπειών και συναφών ισχυρισμών θα περάσουμε αμέσως τώρα ξεκινώντας από την περιοχή της φιλοσοφίας των μαθηματικών.

4. Φιλοσοφία των μαθηματικών

Τα θεωρήματα μη πληρότητας του Gödel έχουν αξιοποιηθεί πολλαπλώς στη φιλοσοφία των μαθηματικών. Η πιο γνωστή εφαρμογή, και εκείνη που κατά κοινή πεποίθηση έδωσε ένα οριστικό αποτέλεσμα, είναι η υπονόμευση του φιλόδοξου προγράμματος για τη θεμελίωση των μαθηματικών που είχε προτείνει ο David Hilbert.[17]

Το πρόγραμμα είχε σχεδιαστεί ως «απάντηση» στην κατάσταση των μαθηματικών στις αρχές του 20ου αιώνα. Εκείνη την εποχή, τα μαθηματικά βρίσκονταν στις «καλύτερες μέρες τους» και, ταυτόχρονα, στις «χειρότερες μέρες τους».[18] Ήδη από τα τέλη του 19ου αιώνα ο Cantor είχε δημιουργήσει για τους μαθηματικούς έναν «παράδεισο» (κατά την έκφραση του Hilbert [1926] 1993, σ. 150) προσφέροντας μια γοητευτική πραγμάτευση του απείρου. Επιπλέον οι συμβολές των Cauchy, Bolzano και Weierstrass στην πραγματική ανάλυση συγκροτούσαν μια στέρεη θεμελίωση του απειροστικού λογισμού: οι ε-δ ορισμοί των ορίων απέφευγαν αοριστίες όπως «άπειρα μεγάλο», «άπειρα μικρό» ή «απειροστό». Ωστόσο, οι νέες θεωρίες εξακολουθούσαν να βασίζονται σε άπειρες ολότητες (τομές Dedekind, όρια ακολουθιών Cauchy, κ.ά.) Κατά δήλωση του ίδιου του Hilbert ([1926] 1993, σ. 146), η μαθηματική ανάλυση παρέμενε «μια συμφωνία του απείρου».

Από την άλλη πλευρά, όμως, διάφορα παράδοξα και αντινομίες δημιουργούσαν βάσιμες ανησυχίες ότι κάποιες μαθηματικές θεωρίες ήταν ασυνεπείς. Η μαθηματική κοινότητα βίωνε μια αίσθηση κρίσης θεμελίων. Και πολλοί μαθηματικοί και φιλόσοφοι (ο Poincaré, ο Russell, οι ιντουισιονιστές, κ.ά.) είχαν προτείνει δραστικούς περιορισμούς της υπάρχουσας μαθηματικής πρακτικής (όπως η απαγόρευση των μη κατηγορηματικών ορισμών, η θεωρία των τύπων, ο εξοβελισμός του πραγματικού απείρου ή η τροποποίηση της κλασικής λογικής).

Αφετηρία του Hilbert ήταν η θέση ότι υπάρχει μια περιοχή των μαθηματικών, η *περατοκρατική θεωρία αριθμών* (στο εξής, ΠΘΑ), στην οποία οι τυχόν εσφαλμένες πεποιθήσεις μας δεν μπορούν παρά να οφείλονται σε «απροσεξία». Γενικότερα, ο Hilbert και οι συνεργάτες του υποστήριζαν ότι το αντικείμενο της ΠΘΑ είναι *ουσιώδες για όλη την ανθρώπινη σκέψη*:

> Προϋπόθεση για να χρησιμοποιήσουμε λογικές συναγωγές και να εκτελέσουμε λογικές πράξεις είναι να έχει ήδη δοθεί κάτι σαν παράσταση: δηλαδή, συγκεκριμένα εξω-λογικά αντικείμενα δοσμένα στην εποπτεία ως άμεσες εμπειρίες πριν από κάθε σκέψη. Για να είναι ασφαλής η λογική συναγωγή, πρέπει τα αντικείμενα αυτά να μπορούν να εποπτευτούν από κάθε τους πλευρά, και το γεγονός ότι παρουσιάζονται, ότι διαφέρουν μεταξύ τους, ότι το ένα ακολουθεί το άλλο, ή ότι είναι συνδυασμένα μεταξύ τους, πρέπει να δίδεται άμεσα στην εποπτεία μαζί με τα αντικείμενα, ως κάτι που δεν επιδέχεται παραπέρα αναγωγή σε κάτι άλλο ή δεν χρειάζεται αναγωγή. Αυτή είναι η βασική φιλοσοφική θέση που θεωρώ αναγκαία όχι μόνο για τα μαθηματικά αλλά, γενικότερα, για κάθε επιστημονική σκέψη, κατανόηση και επικοινωνία. (Hilbert [1926] 1993, σ. 151)

Η βασική ιδέα φαίνεται να είναι ότι προϋπόθεση κάθε σκέψης ή συλλογισμού είναι η δυνατότητα χρήσης συμβόλων με κάποιους κανόνες. Έτσι η ΠΘΑ, ακόμη κι αν δεν είναι απολύτως βέβαιη, είναι όσο βέβαιη επιτρέπουν τα ανθρώπινα μέτρα. Δεν υπάρχει προνομιακότερη, γνωσιολογικά πιο σίγουρη, περιοχή σκέψης από την ΠΘΑ. Γι’ αυτό ο Hilbert ταυτίζει την ΠΘΑ με τα *πραγματικά* μαθηματικά ενώ εντάσσει όλες τις υπόλοιπες μαθηματικές θεωρίες (τα «μαθηματικά του απείρου»: πραγματική και μιγαδική ανάλυση, θεωρία συνόλων, κ.λπ.) στην περιοχή των *ιδεατών* μαθηματικών.

Πάντως, ούτε ο ίδιος ο Hilbert ούτε οι συνεργάτες του προσδιόρισαν με απόλυτη διαύγεια και σαφήνεια το *περιεχόμενο* της ΠΘΑ: ποια αντικείμενα μελετά η ΠΘΑ, ποιες προτάσεις για τέτοια αντικείμενα έχουν νόημα κατά την ΠΘΑ, ποιες πράξεις για τέτοια αντικείμενα είναι νόμιμες κατά την ΠΘΑ και ποιες αποδεικτικές μέθοδοι για τέτοιες προτάσεις είναι αποδεκτές κατά την ΠΘΑ.[19] Η πλέον δημοφιλής ερμηνεία ταυτίζει την ΠΘΑ με το σύστημα για τους φυσικούς αριθμούς που έχει γίνει γνωστό ως *βασική (στοιχειώδης) αναδρομική αριθμητική*. Όμως η υπαρκτή ασάφεια όσον αφορά την οριοθέτηση των περατοκρατικών μέσων δημιουργεί μια ασάφεια όσον αφορά τους στόχους του προγράμματος του Hilbert και, κατ’ ακολουθία, μια ασάφεια όσον αφορά τις επιπτώσεις που είχαν τα θεωρήματα μη πληρότητας του Gödel για το πρόγραμμα αυτό.

Ωστόσο, σύμφωνα με τον Hilbert, η διάκριση μεταξύ πραγματικών και ιδεατών μαθηματικών είναι εξαιρετικά σημαντική. Τα πραγματικά μαθηματικά δεν είναι ένα παιχνίδι χωρίς νόημα ή ένα σύνολο συναγωγών από ανερμήνευτα αξιώματα: οι προτάσεις της ΠΘΑ έχουν νόημα και αντικείμενο που δίνεται σε ένα είδος πρωτογενούς εποπτείας, περίπου *καντιανής* φύσης. Αντίθετα, τα ιδεατά μαθηματικά πρέπει να αντιμετωπίζονται ως ένα *φορμαλιστικό* παιχνίδι με σαφώς διατυπωμένους κανόνες, σαν τους κανόνες σχηματισμού και μετασχηματισμού που παρέχει η τυπική λογική.[20]

Έτσι μια θεωρία των ιδεατών μαθηματικών δεν έχει αξιώσεις αλήθειας αλλά μόνο συνέπειας. Αλλά για να είμαστε κατά το δυνατόν βέβαιοι για τη συνέπεια μιας τέτοιας θεωρίας, η απόδειξη της συνέπειάς της πρέπει να πραγματοποιηθεί χρησιμοποιώντας μόνον περατοκρατικά μέσα. Και τούτο αποτελεί τον πυρήνα του προγράμματος του Hilbert: *τυποποίηση κάθε μαθηματικής θεωρίας σε αξιωματική μορφή και απόδειξη με περατοκρατικά μέσα της συνέπειάς της*. Το κλειδί για την πραγμάτωση του προγράμματος το έδινε η τεχνική της αριθμητικοποίησης της σύνταξης της γλώσσας μιας θεωρίας. Αυτή η τεχνική μας επιτρέπει, όπως είδαμε στην ενότητα 1, να εκφράζουμε προτάσεις για μεταμαθηματικές ιδιότητες θεωριών, όπως η συνέπεια, ως προτάσεις για φυσικούς αριθμούς.

Ποιες, λοιπόν, ήταν οι συνέπειες των θεωρημάτων μη πληρότητας του Gödel για το πρόγραμμα του Hilbert; Η επικρατούσα άποψη είναι ότι εξαιτίας των εν λόγω θεωρημάτων το πρόγραμμα οδηγήθηκε σε οριστική χρεοκοπία. Και το κύριο επιχείρημα υπέρ της άποψης αυτής διατυπώνεται συνήθως ως εξής: το πρόγραμμα του Hilbert απαιτούσε την απόδειξη της συνέπειας της αριθμητικής με τη βοήθεια μόνον της αριθμητικής ενώ το 2οΘΜΠ του Gödel έδειξε ακριβώς ότι μια τέτοια απόδειξη δεν είναι δυνατή.[21]

Όμως η κατάσταση είναι λίγο πιο σύνθετη. Πράγματι, το πρόγραμμα του Hilbert απαιτεί μια περατοκρατική απόδειξη της συνέπειας της PA. Και το 2οΘΜΠ συνεπάγεται ότι η πρόταση $Cons(\mathrm{PA})$ δεν είναι αποδείξιμη στην PA, πόσο μάλλον στο περατοκρατικό «κομμάτι» της PA. Το ίδιο ισχύει για οποιαδήποτε άλλη τυπική θεωρία T που επεκτείνει την PA. Το πρόγραμμα του Hilbert απαιτεί μια περατοκρατική απόδειξη της συνέπειας της T. Αλλά, *εφόσον* αυτή η απόδειξη πρέπει (α) να είναι τυποποιήσιμη εντός της PA και, κατά μείζονα λόγο, εντός της T και (β) να συνίσταται στη τυπική συναγωγή της πρότασης $Cons(T)$ από την T, μια τέτοια απόδειξη δεν είναι δυνατή, αν η T είναι συνεπής, σύμφωνα με το 2οΘΜΠ.

Παρατηρήστε ότι η οριστική καταδίκη του προγράμματος του Hilbert στο φως του 2ουΘΜΠ εξαρτάται από δυο παραδοχές. Σύμφωνα με την πρώτη, κάθε περατοκρατική μέθοδος απόδειξης μπορεί να τυποποιηθεί στην PA. Σύμφωνα με τη δεύτερη, ο ισχυρισμός ότι μια θεωρία T, του είδους που μας ενδιαφέρει, είναι συνεπής εκφράζεται ακριβώς από την πρόταση $Cons(T)$ του Gödel. Και οι δυο αυτές παραδοχές, μολονότι εύλογες από πρώτη άποψη, έχουν αμφισβητηθεί.[22] Εφαλτήριο για την αμφισβήτηση της πρώτης παραδοχής αποτελεί η ασάφεια της οριοθέτησης των περατοκρατικών μέσων που επισημάναμε παραπάνω. Αυτή η ασάφεια επιτρέπει κατ' αρχήν τη δυνατότητα να υπάρχουν αποδεκτοί από περατοκρατική σκοπιά κανόνες που δεν τυποποιούνται στην PA ή ακόμη και τη δυνατότητα η ΠΘΑ να είναι εγγενώς άτυπη. Η δεύτερη παραδοχή έχει αμφισβητηθεί στη βάση του ενδεχομένου να υπάρχουν συμβολικές εκφράσεις της συνέπειας μιας θεωρίας T, του είδους που μας ενδιαφέρει, οι οποίες διαφέρουν από την $Cons(T)$ και είναι ελεύθερες κωλυμάτων παρόμοιων με αυτά που ορθώνει το 2οΘΜΠ.

Δεν θα επιχειρήσω εδώ μια αξιολόγηση των συναφών επιχειρημάτων και θέσεων. Εκείνο που ενδιαφέρει κυρίως στην κατεύθυνση της σκιαγράφησης μιας διαχωριστικής γραμμής ανάμεσα στην κατανόηση και την παρανόηση των θεωρημάτων μη πληρότητας είναι το εξής. Το συμπέρασμα που επιβάλλει *το ίδιο* το 2οΘΜΠ και που μπορεί να αξιοποιηθεί ως προκείμενη σε ένα επιχείρημα για τη βιωσιμότητα του προγράμματος του Hilbert δεν μπορεί παρά να έχει τη μορφή *υποθετικής* πρότασης:

(8) Αν κάθε περατοκρατική μέθοδος απόδειξης μπορεί να τυποποιηθεί στην PA και αν η συνέπεια μιας θεωρίας T εκφράζεται από την πρόταση $Cons(T)$, τότε η συνέπεια οποιασδήποτε συνεπούς θεωρίας που επεκτείνει την PA δεν μπορεί να αποδειχθεί με περατοκρατικά μέσα.

Όπως θα προσπαθήσω να επιχειρηματολογήσω στο υπόλοιπο κείμενο αυτή είναι η λογική μορφή κάθε συμπεράσματος που μπορεί να εξαχθεί από ένα μαθηματικό θεώρημα υπέρ ή κατά κάποιας φιλοσοφικής θέσης. Και συχνά οι παρανοήσεις πηγάζουν από ηθελημένη ή αθέλητη παράβλεψη αυτού του, μάλλον τετριμμένου, διδάγματος. Αλλά ας περάσουμε τώρα σε μια σύντομη αναφορά της επίδρασης των θεωρημάτων μη πληρότητας στην περιοχή της φιλοσοφίας των φυσικών επιστημών.

5. Φιλοσοφία των φυσικών επιστημών

Τουλάχιστον από την εποχή του Νεύτωνα, οι φυσικές θεωρίες που περιγράφουν με επιτυχία τον κόσμο της εμπειρίας είναι διατυπωμένες σε μαθηματική γλώσσα. Το γεγονός αυτό θέτει μια πρόκληση στη φιλοσοφία. Πώς είναι δυνατόν τα μαθηματικά να είναι *a priori* –δηλαδή, να είναι γνώσιμα ανεξάρτητα από συγκεκριμένες εμπειρίες (πειράματα ή παρατηρήσεις)– και, ταυτόχρονα, να εφαρμόζονται με ακρίβεια στη μελέτη του κόσμου της εμπειρίας;

Το πρόβλημα είναι ιδιαίτερα οξύ για τους *εμπειριστές* σύμφωνα με τους οποίους κάθε γνώση γεγονότων της πραγματικότητας –γνώση που ορισμένοι διακρίνουν από τη γνώση αμιγώς λογικών σχέσεων μεταξύ εννοιών– «προέρχεται» τελικώς από την εμπειρία. Το κίνημα των *λογικών θετικιστών* ή *λογικών εμπειριστών*[23] που αναδύθηκε στη γερμανόφωνη Ευρώπη κατά τις πρώτες δεκαετίες του 20ου αιώνα επιχείρησε να αντιμετωπίσει εκ νέου αυτή την πρόκληση. Η γενική ιδέα, η προέλευση της οποίας ανάγεται στη φιλοσοφία του Hume και στον λογικισμό, είχε ως εξής: *η λογικο-μαθηματική γνώση είναι a priori επειδή ακριβώς είναι γνώση της χρήσης ενός γλωσσικού πλαισίου*. Προκειμένου να περιγράψουμε τον φυσικό κόσμο, επιλέγουμε κατ' αρχήν *συμβατικά* ένα γλωσσικό πλαίσιο, μέρος του οποίου είναι οι λογικο-μαθηματικές θεωρίες. Οι προτάσεις της λογικής και των μαθηματικών εντός ενός τέτοιου πλαισίου δεν μπορούν παρά να εφαρμόζονται στον κόσμο της εμπειρίας επειδή ακριβώς η αλήθειά τους είναι ανεξάρτητη από το πώς έχει ο κόσμος της εμπειρίας. Και τούτο γιατί, από *σημασιολογική* άποψη, οι εν λόγω προτάσεις είναι ταυτολογικές ή αναλυτικές – δηλαδή, γιατί η αλήθειά τους προκύπτει μόνον από τη λογική δομή τους και το νόημα των «όρων» που περιέχουν. Η δε χρησιμότητα της λογικής και των μαθηματικών στη φυσική επιστήμη έγκειται στο ότι επιτρέπουν τον μετασχηματισμό εμπειρικών προτάσεων σε άλλες εμπειρικές προτάσεις με κατ' ανάγκην διατήρηση της αλήθειας.[24]

Σύμφωνα με αυτή την προσέγγιση της διάκρισης λογικο-μαθηματικής και εμπειρικής γνώσης, η λογική και τα μαθηματικά αποτελούν ζήτημα συμβατικών επιλογών. Διακεκριμένος υποστηρικτής αυτής της προσέγγισης ήταν ο Carnap, ο οποίος διακήρυσσε ότι για τη συμβατική επιλογή λογικο-μαθηματικού πλαισίου ίσχυε η ακόλουθη *αρχή ανεκτικότητας*:

> Στη λογική δεν υπάρχουν ηθικοί φραγμοί. Καθένας είναι ελεύθερος να οικοδομήσει τη δική του λογική –δηλαδή, τη δική του μορφή γλώσσας– όπως επιθυμεί…. Η λογική, μόλις τυποποιηθεί με ακριβή τρόπο, φαίνεται ότι δεν είναι παρά η σύνταξη είτε κάποιας συγκεκριμένης γλώσσας είτε των γλωσσών γενικώς. (Carnap [1934] 1937, § 17, 62)

Ακόμη και σε μεταγενέστερα κείμενα, ο Carnap επέμενε στη συμβασιοκρατική θέση ότι η επιλογή ενός γλωσσικού πλαισίου για την επιστήμη κρίνεται μόνον εκ των υστέρων ως επιτυχής ή μη ανάλογα με το πόσο εξυπηρετεί τους στόχους της έρευνας:

> Η αποδοχή ή απόρριψη των αφηρημένων γλωσσικών μορφών, ακριβώς όπως η αποδοχή ή απόρριψη οποιωνδήποτε γλωσσικών μορφών σε οποιαδήποτε περιοχή της επιστήμης, θα κριθεί τελικά από την αποτελεσματικότητα των εργαλείων της, από το λόγο των αποτελεσμάτων προς το πόσο και τη συνθετότητα των προσπαθειών που καταβλήθηκαν…. *Ας είμαστε προσεκτικοί όταν κάνουμε βεβαιώσεις και κριτικοί όταν τις εξετάζουμε, αλλά ας είμαστε ανεκτικοί όσον αφορά τις γλωσσικές μορφές που επιτρέπουμε.* (Carnap [1950] 1988, σ. 61)

Το «σύστημα των αριθμών» που σκιαγράφησε ο Carnap ([1950] 1988, σ. 37-39) είναι ακριβώς ένα παράδειγμα τέτοιου γλωσσικού πλαισίου για την επιστήμη.

Ο Gödel, όντας ο ίδιος μέλος του Κύκλου της Βιέννης, βρισκόταν σε στενή επαφή με τους λογικούς θετικιστές. Όμως οι πλατωνιστικές του προδιαθέσεις απέναντι στα μαθηματικά δεν του επέτρεπαν να υιοθετήσει θέσεις παρόμοιες με εκείνες του Carnap. Σε μια εργασία που είχε υποσχεθεί να υποβάλει προκειμένου να συμπεριληφθεί σε ένα τόμο για τη φιλοσοφία του Carnap, αλλά την οποία δεν υπέβαλε γιατί δεν ήταν ικανοποιημένος με το αποτέλεσμα αφού την είχε γράψει ήδη έξι φορές, ο Gödel ([1953/9] 1995) ανέπτυξε ένα επιχείρημα ενάντια στις απόψεις του Carnap που μπορεί να ανασυγκροτηθεί ως εξής.[25]

Ας θεωρήσουμε ένα *γκεντελιανό* σύστημα λογικο-μαθηματικών κανόνων S.[26] Το S μπορεί να αποκληθεί *συντακτικό* μόνον αν *δεν* συνεπάγεται την αλήθεια ή το ψεύδος οποιασδήποτε εμπειρικής πρότασης (αφού η αληθοτιμή μιας *εμπειρικής* πρότασης δεν μπορεί να είναι *αποκλειστικά* θέμα *σύνταξης*). Αλλά αυτή η απαίτηση μπορεί να ικανοποιηθεί μόνον αν το S είναι *συνεπές* (γιατί, αλλιώς, θα συνεπαγόταν *κάθε* πρόταση). Έπεται, με τη βοήθεια του 2ουΘΜΠ, ότι η συνέπεια και ο συντακτικός χαρακτήρας του S μπορεί να δικαιολογηθεί μόνο με επίκληση λογικο-μαθηματικών προτάσεων και κανόνων *έξω* από το σύστημα S. Επομένως, η λογική και τα μαθηματικά δεν είναι αποκλειστικά ζήτημα επιλογής συντακτικών κανόνων.

Αντιμέτωπος με αυτό το επιχείρημα, θα μπορούσε κανείς να ρωτήσει: γιατί, όμως, να θέλουμε να *δικαιολογήσουμε* τη συνέπεια και το συντακτικό χαρακτήρα ενός τέτοιου συστήματος S; Η απάντηση σχετίζεται με τη χρησιμότητα του συστήματος στη μελέτη του εμπειρικού κόσμου. Το S *πρέπει* να είναι συνεπές αν ο ρόλος των λογικο-μαθηματικών κανόνων που το απαρτίζουν είναι να επιτρέπουν τη συναγωγή νέων εμπειρικών προτάσεων από δεδομένες εμπειρικές προτάσεις *με κατ' ανάγκην διατήρηση της αλήθειας*. Και εμείς πρέπει να *γνωρίζουμε* ότι το S είναι συνεπές προκείμενου να μπορέσουμε να *εξηγήσουμε* γιατί το S, μολονότι συμβατικά επιλεγμένο, εφαρμόζεται με ακρίβεια στη μελέτη του κόσμου της εμπειρίας. Το γενικό συμπέρασμα έχει πάλι τη μορφή *υποθετικής* πρότασης: η αποδοχή της θέσης ότι οι λογικο-μαθηματικοί κανόνες είναι συντακτικές (γλωσσικές) συμβάσεις συνεπάγεται την αδυναμία εξήγησης της απροσδόκητης εφαρμοσιμότητας των μαθηματικών στις εμπειρικές επιστήμες. Μάλιστα, μπορεί κανείς να διαγνώσει «πίσω» από το συμπέρασμα την ύπαρξη ενός πλήθους πρόσθετων υπόρρητων, μεθοδολογικών αλλά και μεταφυσικών, υποθέσεων: ότι η εξήγηση της εφαρμοσιμότητας των μαθηματικών στις εμπειρικές επιστήμες πρέπει να ακολουθεί το υπόδειγμα αυστηρότητας της τυπικής λογικής, ότι η απόδειξη της συνέπειας κάθε τυπικού συστήματος S συνίσταται στην απόδειξη της αντίστοιχης πρότασης *Cons*(S) την αποδειξιμότητα της οποίας διερευνά το 2οΘΜΠ, ότι ο εμπειρικός κόσμος είναι ανεξάρτητος από τις λογικο-μαθηματικές μας πρακτικές, κ.λπ. Δεν προτίθεμαι να ισχυριστώ ότι οι εν λόγω υποθέσεις δεν είναι εύλογες ή ότι δεν θα ήταν αποδεκτές από λογικούς θετικιστές. Στόχος, για μια ακόμη φορά, είναι να τονιστεί η μορφή του συμπεράσματος που μπορεί να εξαχθεί και η προσοχή με την οποία πρέπει να επιχειρήσει κανείς την εξαγωγή του.

6. Φιλοσοφία του νου

Η *μηχανοκρατία* της νέας επιστήμης του 17ου αιώνα έθεσε επιτακτικά στη φιλοσοφία το ερώτημα εάν όλες οι λειτουργίες του ανθρώπινου νου μπορούν να εκτελεστούν από μηχανές. Αν τα πάντα είναι ύλη σε κίνηση και επιδέχονται αυστηρή περιγραφή με μαθηματικές έννοιες, τι εμποδίζει την προσομοίωση της ανθρώπινης νόησης σε μηχανές; Ένας φιλόσοφος που απάντησε «*Τίποτα!*» ήταν ο Thomas Hobbes:

> Συλλογισμός [reason], με αυτή τη σημασία, δεν είναι παρά υπολογισμός [reckoning] (δηλαδή, πρόσθεση και αφαίρεση) των συνεπειών των γενικών ονομάτων που έχουν συμφωνηθεί, για τη διάκριση και σήμανση των σκέψεων μας. (Hobbes [1651] 1985, σ. 111)

Για αυτό το λόγο ο Hobbes μπορεί κάλλιστα να θεωρηθεί πρόδρομος σύγχρονης *υπολογιστικής θεωρίας του νου*. Πρόκειται για εκείνη την προσέγγιση στη φιλοσοφία του νου σύμφωνα με την οποία κάθε σκέψη (τουλάχιστον, όσον αφορά προτασιακά περιεχόμενα) είναι ένας υπολογισμός που θα μπορούσε να εκτελεστεί από μια μηχανή.[27]

Όμως τα θεωρήματα μη πληρότητας του Gödel φαίνεται να αναδεικνύουν όρια στο τι μπορούν να επιτύχουν κάποιες μηχανές. Έτσι από τις αρχές της δεκαετίας του 1960 ξεκίνησε μια έντονη συζήτηση για τις συνέπειες των θεωρημάτων μη πληρότητας όσον αφορά τη δυνατότητα «μηχανοποίησης» της ανθρώπινης σκέψης. Κατά αυτής της δυνατότητας έχουν ταχθεί ο John Lucas (1961), ο Roger Penrose (1989, 1994) και, πιο πρόσφατα, ο Storrs McCall (1999, 2001). Η συναφής βιβλιογραφία είναι ιδιαίτερα εκτεταμένη. Και η συζήτηση συνεχίζεται. Εδώ θα περιοριστώ μόνο στα πρώτα στάδια της συζήτησης.

Το αρχικό επιχείρημα του Lucas κατά της υπολογιστικής σύλληψης του ανθρώπινου νου ήταν, σε δικά του λόγια, το εξής:

> Το θεώρημα του Gödel πρέπει να εφαρμόζεται σε μηχανές της κυβερνητικής, γιατί είναι ουσιώδες χαρακτηριστικό μιας μηχανής το να αποτελεί συγκεκριμένη πραγμάτωση ενός τυπικού συστήματος. Έπεται ότι δεδομένης οποιασδήποτε μηχανής που είναι συνεπής και επαρκής για να κάνει απλή αριθμητική, υπάρχει ένας τύπος τον οποίο αδυνατεί να παραγάγει ως αληθή – ο τύπος που είναι μη αποδείξιμος στο σύστημα – αλλά του οποίου *την αλήθεια μπορούμε εμείς να δούμε*. Έπεται ότι καμία μηχανή δεν μπορεί να αποτελέσει πλήρες και επαρκές μοντέλο του νου, ότι οι νόες είναι ουσιωδώς διαφορετικοί από μηχανές. (Lucas 1961, σ. 113, δική μου έμφαση)

Από το 1961 μέχρι σήμερα, η ίδια βασική ιδέα έχει αποτελέσει τον πυρήνα πολλών, ίσως πιο εκλεπτυσμένων, επιχειρημάτων. Η δε κριτική των αντιπάλων επικεντρώνεται σε ένα σημείο: *Μπορούμε εμείς να «δούμε» ότι η πρόταση* $Gödel, G_T$*, ενός γκεντελιανού τυπικού συστήματος* T *είναι αληθής;* Έχουμε λόγους να πιστεύουμε ότι η G_T είναι αληθής ακριβώς στον βαθμό που έχουμε λόγους να πιστεύουμε ότι το T είναι συνεπές. Αλλά μπορούμε εμείς να «δούμε» ότι το T είναι συνεπές; Εξαρτάται. Αν «βλέπω» σημαίνει «αποδεικνύω τυπικά», τότε το 2οΘΜΠ επιβάλλει περιορισμούς στη δυνατότητά μας να «δούμε» τη συνέπεια ενός γκεντελιανού τυπικού συστήματος. Ο Lucas είχε επίγνωση αυτής της περιπλοκής. Για αυτό, μερικές σελίδες παρακάτω στο ίδιο άρθρο, έγραφε:

> Το μόνο που έδειξε ο Gödel είναι ότι ένας νους δεν μπορεί να παραγάγει μια τυπική απόδειξη της συνέπειας ενός τυπικού συστήματος μέσα στο ίδιο το σύστημα· αλλά δεν υπάρχει αντίρρηση στο να βγει έξω από το σύστημα και στο να παραγάγει άτυπα επιχειρήματα για τη συνέπεια είτε ενός τυπικού συστήματος είτε κάποιου άλλου συστήματος λιγότερο τυπικού και λιγότερο συστηματοποιημένου.… Μου φαίνεται, επομένως, σωστό και εύλογο ένας νους να ισχυρίζεται τη δική του συνέπεια … (Lucas 1961, σ. 124-125)

Όμως η ένσταση από τους «συνηγόρους των δυνατοτήτων» των μηχανών ήταν άμεση:[28]

> Αλλά είναι σαφές ότι δεν μπορεί να κάνει το ίδιο η μηχανή Maud; Η Maud είναι περιορισμένη σε εκείνα τα πράγματα για τα οποία μπορεί να προσφέρει τυπικές$_{\text{Maud}}$ αποδείξεις. Αλλά περιορίζει αυτό τις άτυπες (δηλαδή, τις μη τυποποιήσιμες στη Maud) αποδείξεις που μπορεί να επινοήσει; Δεν είναι σαφές ότι τις περιορίζει. Και τούτο διότι η Maud μπορεί να εφαρμόσει το επιχείρημα του Gödel στον εαυτό της: από το 2οΘΜΠ μπορεί να αποδείξει την '$Cons(\text{Maud}) \to G_{\text{Maud}}$'. Δεν θα μπορούσε, λοιπόν, έτσι να «πείσει τον εαυτό της ότι η G_{Maud}, μολονότι μη αποδείξιμη-στη-Maud, ήταν εντούτοις –στην πραγματικότητα, για αυτόν ακριβώς τον λόγο– αληθής»; (Benacerraf 1967, σ. 20)

Βέβαια θα μπορούσε κανείς να αντιτείνει ότι καμία μηχανή δεν μπορεί να *πείσει τον εαυτό της* για την αλήθεια ενός τύπου. Εξάλλου τι ακριβώς θα σήμαινε αυτό; Το πολύ που μπορεί να κάνει μια μηχανή είναι να παραγάγει κάποιον τύπο στην έξοδο. Αλλά, όπως ορθά παρατηρεί ο Benacerraf (1967, σ. 20), αν *αυτός* είναι ο λόγος για τον οποίο μια μηχανή δεν μπορεί να κάνει ό,τι μπορεί να κάνει ο ανθρώπινος νους, τότε η επίκληση των ΘΜΠ του Gödel είναι περιττή.

Αδικώντας κατάφωρα την πολυπλοκότητα της συζήτησης που ξεκίνησαν οι Lucas και Benacerraf να σταθώ εδώ για να πω το εξής. Αν δεν είναι σαφές με *ποιες* μεθόδους μπορώ να πείσω τον εαυτό και για *ποια* πράγματα, τότε μπορώ να πείσω τον εαυτό μου ότι είναι ότι κάποιες προτάσεις είναι αληθείς επειδή ακριβώς τις πιστεύω. Να ένα παράδειγμα:

(9)
1. Πιστεύω ότι έχω τουλάχιστον μια ψευδή πεποίθηση.
2. Η *παραπάνω* πεποίθησή μου είναι ή αληθής ή ψευδής.
3. Αν η *παραπάνω* πεποίθησή μου είναι αληθής, τότε έχω τουλάχιστον μια ψευδή πεποίθηση.
4. Αν η *παραπάνω* πεποίθησή μου είναι ψευδής, τότε έχω τουλάχιστον μια ψευδή πεποίθηση.

∴ Έχω τουλάχιστον μια ψευδή πεποίθηση.

Το παράδειγμα είναι κατάλληλο, όχι μόνον λόγω της αυτο-αναφορικής του διάστασης, αλλά και για ένα επιπλέον λόγο. Θα μπορούσε να αξιοποιηθεί στην απόδειξη του ισχυρισμού ‘Δεν είμαι τέλειος’ προκειμένου να στηρίξει το επιχείρημα για την ύπαρξη του Θεού που αναπτύσσει ο Descartes στον *Τρίτο Στοχασμό*. Έτσι θα φθάναμε εγγύτερα στη «μαθηματική θεολογία», για να χρησιμοποιήσω μια έκφραση του Benacerraf (1967, σ. 10) – σε στοχασμούς που εντάσσονται σε αυτό που ονομάζω παρακάτω «φιλοσοφία … πολύ γενικότερη».

Αλλά προτού ασχοληθούμε με τέτοιους στοχασμούς, ας περιγράψω επιγραμματικά τρεις ομάδες προβλημάτων θα έπρεπε να αντιμετωπιστούν προκειμένου να αποφανθούμε αν τα ΘΜΠ έχουν τελεσίδικες συνέπειες όσον αφορά τη δυνατότητα «μηχανοποίησης» της ανθρώπινης νόησης.

Η πρώτη ομάδα προβλημάτων αφορά το είδος των μη αποφασίσιμων προτάσεων σε γκεντελιανά συστήματα που προτείνονται ως μοντέλα μηχανών. Αν M είναι μια μηχανή, ποια ακριβώς μη-αποφασίσιμη-στο-σύστημα-της-M πρόταση προκύπτει; Με ποιο τρόπο; Και υπό ποιες προϋποθέσεις; Η απάντηση τέτοιων ερωτημάτων απαιτεί ασφαλή κατανόηση, όχι μόνον του περιεχομένου των ΘΜΠ, αλλά και των αποδείξεών τους. Για τεκμηρίωση αυτού του ισχυρισμού, ο αναγνώστης καλείται να μελετήσει την κριτική που άσκησε ο Tennant (2001) στο επιχείρημα με το οποίο ο McCall (1999) προσπάθησε να δικαιολογήσει τη θέση ότι οι δυνατότητες του ανθρώπινου νου υπερβαίνουν εκείνες των μηχανών Turing. Το αδύνατο σημείο του επιχειρήματος του McCall συνίσταται σε μια σύγχυση ανάμεσα σε μη αποφασίσιμες προτάσεις *τύπου Gödel* (που προϋποθέτουν ω-συνέπεια του συστήματος) και μη αποφασίσιμες προτάσεις *τύπου Rosser* (για τις οποίες αρκεί η «απλή» συνέπεια του συστήματος).

Η δεύτερη ομάδα προβλημάτων αφορά στο τι μπορεί να παραγάγει στην έξοδο μια μηχανή. Μπορεί, λόγου χάριν, μια μηχανή Μ να «διαπιστώσει» ότι η Μ-αποδειξιμότητα και η αλήθεια δεν συμπίπτουν; Και πώς; Εδώ μπορούν να αξιοποιηθούν μαθηματικά αποτελέσματα *εφόσον* η μηχανή έχει οριστεί ως μαθηματικό αντικείμενο. Αλλά και αν έχει έτσι οριστεί, είναι ο ορισμός αρκετά ευρύς; Μεταξύ άλλων, ισχύει η θέση του Church;

Η τρίτη ομάδα προβλημάτων αφορά στο τι μπορεί να «δει», «πιστέψει», «γνωρίσει», κ.λπ., ένας άνθρωπος. Και για τέτοια ζητήματα δεν μπορούν να υπάρξουν θεωρήματα της λογικής ή των μαθηματικών ανεξάρτητα από ουσιώδεις φιλοσοφικές θέσεις για την ανθρώπινη φύση και νόηση και υποθέσεις για την λογικο-μαθηματική διατύπωση αυτών των θέσεων. Το δίδαγμα, ξανά, είναι ότι το οποιοδήποτε κατηγορηματικό συμπέρασμα δεν μπορεί παρά να λάβει τη μορφή υποθετικής πρότασης

…

ω. Φιλοσοφία … πολύ γενικότερη! Επίλογος

Μερικοί επικαλούνται τα ΘΜΠ για να υποστηρίξουν ισχυρισμούς που είναι πολύ πιο εντυπωσιακοί από αυτούς που προσπάθησα να σκιαγραφήσω στα προηγούμενα. Για παράδειγμα, ο Paul Davies ανακαλύπτει στα ΘΜΠ ίχνη της οδού προς το μυστικισμό:

> Είμαστε αποκλεισμένοι από την έσχατη γνώση, την έσχατη εξήγηση, από τους ίδιους τους κανόνες συλλογισμού που μας παρακίνησαν αρχικά να αναζητήσουμε μια τέτοια εξήγηση. Αν επιθυμούμε να προχωρήσουμε παραπέρα, πρέπει να ενστερνιστούμε μια διαφορετική έννοια «κατανόησης» από εκείνη της ορθολογικής εξήγησης. Είναι δυνατόν να είναι το μυστικιστικό μονοπάτι ένας δρόμος προς μια τέτοια κατανόηση. Ίσως οι μυστικιστικές ενοράσεις παρέχουν τη μόνη οδό πέρα από τα όρια στα οποία μπορούν να μας οδηγήσουν η επιστήμη και η φιλοσοφία, το μόνο δυνατό μονοπάτι προς το Έσχατο.[29]

Είναι δύσκολο να αποτιμήσει κανείς με ακρίβεια την ορθότητα τέτοιων ισχυρισμών και συλλογισμών, κυρίως λόγω της αοριστίας τους. Και είναι αντιδεοντολογικό να διατυπώνει κανείς κρίσεις για θέσεις άλλων χωρίς να λαμβάνει υπόψη το όλο πλαίσιο. Παρ’ όλα αυτά, ας μου επιτραπεί να χρησιμοποιήσω το παραπάνω απόσπασμα για να επισημάνω δυο παρανοήσεις που ενδέχεται να είναι υπεύθυνες για την εξαγωγή τέτοιων συμπερασμάτων.

Πρώτον, από τα ΘΜΠ του Gödel δεν προκύπτει ότι υπάρχουν *απολύτως* μη αποδείξιμες προτάσεις. Μια πρόταση που δεν αποδεικνύεται στο πλαίσιο ενός τυπικού συστήματος μπορεί να αποδεικνύεται στο πλαίσιο ενός ισχυροτέρου. Ο ίδιος ο Gödel ([1964] 1993, σ. 187-188), σχολιάζοντας τη μη αποφασισιμότητα της εικασίας του συνεχούς του Cantor[30] στην αξιωματική θεωρία συνόλων Zermelo-Fraenkel με αξίωμα επιλογής, ισχυρίστηκε ότι το πρόβλημα ενδέχεται να απαντηθεί με την προσθήκη ενός νέου αξιώματος του οποίου η αλήθεια θα γίνει αποδεκτή «επαγωγικά», δηλαδή εξετάζοντας τη γονιμότητά του σε ανεξάρτητα επαληθεύσιμες συνέπειες – όπως περίπου συμβαίνει και στις καλώς θεμελιωμένες φυσικές θεωρίες. Αυτό, κατά δήλωση του ιδίου, σημαίνει ότι το νέο αξίωμα θα είναι «μόνον πιθανό». Προτείνοντας τη γονιμότητα ως επαγωγικό τεκμήριο για την ανακάλυψη μιας πληρέστερης περιγραφής της ανεξάρτητης από εμάς πραγματικότητας των συνόλων, ο πλατωνιστής Gödel αποποιείται συνειδητά τη *μαθηματική βεβαιότητα*.

Αυτό με φέρνει στο δεύτερο σημείο που θέλω να τονίσω. Το ιδεώδες της μαθηματικής βεβαιότητας, στην «αυθόρμητη φιλοσοφία των μαθηματικών» των περισσοτέρων ανθρώπων, πηγάζει από τη λογικό καθεστώς των μαθηματικών *αποδείξεων*. Μια μαθηματική απόδειξη, εφόσον είναι ένα έγκυρο παραγωγικό επιχείρημα, διατηρεί κατ’ ανάγκην την αλήθεια με την έννοια ότι δεν είναι δυνατόν να είναι οι προκείμενες (τα «αξιώματα») αληθείς αλλά το συμπέρασμα (το «θεώρημα») ψευδές. Όμως, ακριβώς για αυτόν το λόγο, μια μαθηματική απόδειξη δεν προσθέτει νέα γνώση από καθαρά *λογική*, όχι *ψυχολογική*, σκοπιά. Το περιεχόμενο του συμπεράσματος (του «θεωρήματος») *εμπεριέχεται* σε εκείνο του συνόλου των προκειμένων (των «αξιωμάτων»). Επομένως η αύξηση της γνώσης απαιτεί μη παραγωγικές μεθόδους συναγωγής συμπερασμάτων. Αλλά αυτές δεν είναι απαραίτητα «μυστικιστικές». Οι επαγωγικές μέθοδοι μάθησης από την εμπειρία μπορεί να μην εξασφαλίζουν *βεβαιότητα* αλλά είναι ορθολογικές – μολονότι αποτελεί ενδιαφέρον φιλοσοφικό πρόβλημα το *γιατί* είναι. Εξάλλου μια μακρά φιλοσοφική παράδοση έχει καταδείξει ότι το δίπολο «*Βεβαιότητα ή Ανορθολογισμός*» συγκροτεί ψευδή διάζευξη: το ότι είναι λογικώς ή και γνωσιολογικώς δυνατόν να είναι ψευδείς κάποιες πεποιθήσεις μου δεν σημαίνει ότι δεν είμαι ορθολογικώς δικαιολογημένος να έχω αυτές τις πεποιθήσεις. Αν το σφάλμα μπορεί να νοηθεί ως χάσμα ανάμεσα στις πεποιθήσεις μου και σε κάποια «εξωτερική πραγματικότητα», τότε ο μόνος βέβαιος τρόπος να αποφύγω το σφάλμα είναι να μην λέω τίποτε για την «εξωτερική πραγματικότητα».

Οι παραπάνω παρατηρήσεις ίσως φαίνονται τετριμμένες. Σίγουρα θα έπρεπε να είναι τετριμμένες για οποιονδήποτε προπτυχιακό σπουδαστή της φιλοσοφίας. Όμως τα αποτελέσματα του Gödel έχουν τύχει πολύ χειρότερης μεταχείρισης από εκείνη από την οποία τα υπερασπίζονται οι παραπάνω παρατηρήσεις. Ο Jacques Bouveresse ([1999] 2002) έχει αφιερώσει πάνω από ένα κεφάλαιο του βιβλίου του σε ανενδοίαστες «διανοητικές αλχημείες» που έχουν ως «πρώτη ύλη» τα ΘΜΠ.[31] Επιλέγω μόνο μια σχετική δήλωση ως αφετηρία για σχολιασμό και παραπέμπω τον αναγνώστη στο βιβλίο του Bouveresse και τις εκεί αναφορές. Η δήλωση αποδίδεται[32] στον Régis Debray και έχει ως εξής: «Γνωρίζουμε με επιστημονική βεβαιότητα χάρη στο αξίωμα μη-πληρότητας, ότι η “χειραφέτηση του ανθρωπίνου είδους” είναι αυταπάτη, αιώνια και αναγκαία …». Δεν είμαι σε θέση να σχολιάσω αναλυτικά αυτή, ή παρόμοιες με αυτή, δηλώσεις του Debray. Βέβαια, μπορεί κανείς να παρατηρήσει αμέσως την παραπλανητική χρήση του όρου ‘*αξίωμα* μη πληρότητας’ αντί του ‘*θεώρημα* μη πληρότητας’ και να διερωτηθεί κατά πόσον η “χειραφέτηση του ανθρωπίνου είδους” εξαρτάται από την ύπαρξη αλγορίθμων. Θέλω μόνο να σταθώ λίγο στη φράση «Γνωρίζουμε με επιστημονική βεβαιότητα …». Αυτή η φράση αποτελεί συνηθισμένη εισαγωγή σε *επιχειρήματα με επίκληση αυθεντίας*. Τέτοια επιχειρήματα είναι παραπλανητικά εκτός εάν (α) η αυθεντία την οποία επικαλείται το επιχείρημα είναι όντως αυθεντία και μάλιστα στην περιοχή γνώσης στην οποία ανήκει το συμπέρασμα, (β) ο ισχυρισμός της αυθεντίας τον οποίο επικαλείται το επιχείρημα αφορά τη συγκεκριμένη περιοχή γνώσης, και (γ) υπάρχει ομοφωνία ανάμεσα στις αυθεντίες ως προς την αλήθεια του ισχυρισμού αυτού. Ωστόσο στα περισσότερα *φιλοσοφικά* ζητήματα –έστω και αν πρόκειται για φιλοσοφικές συνέπειες μαθηματικών ή επιστημονικών επιτευγμάτων– δεν υπάρχει ομοφωνία μεταξύ ειδικών. Και σε αυτές τις περιπτώσεις καλείται ο καθένας να καταλήξει στα δικά του συμπεράσματα. Αυτό είναι προφανές, παραδείγματος χάριν, σε ζητήματα εφαρμοσμένης ηθικής (ευθανασία, άμβλωση, κατανομή πλούτου, κ.λπ.).

Το να επικαλεστεί κανείς την αυθεντία ενός μαθηματικού θεωρήματος ή ενός επιστημονικού επιτεύγματος για να πείσει κάποιον για την αλήθεια μιας φιλοσοφικής θέσης χωρίς να μπει στον κόπο να αναλύσει με την απαιτούμενη επάρκεια το θεώρημα ή επίτευγμα και χωρίς να αναπτύξει με τη μέγιστη δυνατή σαφήνεια τον τρόπο με τον οποίο αξιοποιεί το θεώρημα ή επίτευγμα, αλλά διεκδικώντας με αυτό τον τρόπο απουσία κριτικής, είναι απλώς απάτη. Και η απάτη είναι θλιβερότερη όταν διαπράττεται από «πολέμιους του επιστημονισμού». Αυτό δεν σημαίνει ότι κανείς δεν δικαιούται να μιλά για τα μαθηματικά ή την επιστήμη αν δεν είναι ειδικός, με τη στενή σημασία, του αντίστοιχου κλάδου. Αντίθετα τα μαθηματικά και η επιστήμη συγκροτούν οργανικό κομμάτι της ευρύτερης πολιτιστικής παραγωγής στην οποία όλοι θα πρέπει να έχουν πρόσβαση. Το αίτημα είναι μάλλον ζήτημα *ηθικής εντιμότητας*. Τα σφάλματα συγχωρούνται, μολονότι πρέπει να διορθώνονται μετά την έκθεση σε εύστοχη κριτική, εφόσον έχει καταβληθεί έντιμη προσπάθεια να αποφευχθούν. Και το τι αποτελεί έντιμη προσπάθεια κάθε φορά εξαρτάται από τον εκάστοτε στόχο. Άλλη προσπάθεια απαιτείται για τη συναγωγή ενός κατηγορηματικού συμπεράσματος, άλλη για την απλή επισήμανση γόνιμων αναλογιών μεταξύ διαφορετικών περιοχών της σκέψης, άλλη για τη δικαιολόγηση ενός απλού εννοιολογικού δανείου, κ.λπ.

Για να δώσω ένα παράδειγμα, παραθέτω ένα απόσπασμα από ένα άρθρο ενός γιαπωνέζου μελετητή της αποδόμησης, του Hiroki Azuma:

> Η αποδόμηση κατά την ανάγνωση γίνεται ανακαλύπτοντας μέσα στο κείμενο ένα σημείο στο οποίο το (γραμματικό) νόημα του επιπέδου-αντικειμένου και το (ρητορικό) νόημα του μετα-επιπέδου θα αναμιγνύονταν αναπόφευκτα το ένα με το άλλο – πράγμα που θα έπρεπε να θεωρηθεί το λογοτεχνικό ανάλογο των μαθηματικών τεχνικών που χρησιμοποίησε ο Gödel στην απόδειξη του θεωρήματός του. (Azuma 2006, σ. 2)

Προσέξτε ότι δηλώνεται ρητά πως πρόκειται για *αναλογία*. Προσέξτε επίσης ότι διαφαίνονται –έστω, *αόριστα*– κάποια σημεία αναλογίας με το υπόβαθρο των θεωρημάτων του Gödel: διάκριση μεταξύ γλώσσας και μεταγλώσσας, αυτοαναφορική «σύγχυση» των δυο επιπέδων, κ.λπ. Επιπλέον ο συγγραφέας χρησιμοποιεί την αναλογία απλώς για να δικαιολογήσει τη χρήση, στο υπόλοιπο του άρθρου, του όρου 'γκεντελιανή αποδόμηση' ('Gödelian deconstruction') με περιεχόμενο που αποδίδει στον Paul de Man.

Θα κλείσω παραθέτοντας τα λόγια με τα οποία ο Kurt Gödel άρχιζε το άρθρο του πάνω στη μη πληρότητα:

> Η ανάπτυξη των μαθηματικών στην κατεύθυνση μεγαλύτερης ακρίβειας έχει οδηγήσει, όπως είναι γνωστό, στην τυποποίηση μεγάλων περιοχών τους, έτσι ώστε να μπορεί κανείς να αποδείξει οποιοδήποτε θεώρημα με τίποτα παραπάνω από λίγους μηχανικούς κανόνες. Τα πλέον περιεκτικά τυπικά συστήματα που έχουν κατασκευαστεί μέχρι τώρα είναι αφενός το σύστημα των *Principia mathematica* και αφετέρου το αξιωματικό σύστημα για τη θεωρία συνόλων των Zermelo-Fraenkel (που αναπτύχθηκε περαιτέρω από τον J. von Neumann). Αυτά τα δυο συστήματα είναι τόσο περιεκτικά ώστε όλες οι μέθοδοι απόδειξης

> που χρησιμοποιούνται σήμερα στα μαθηματικά τυποποιούνται μέσα σε αυτά, δηλαδή ανάγονται σε λίγα αξιώματα και αποδεικτικούς κανόνες. Μπορεί κανείς, λοιπόν, να εικάσει ότι αυτά τα αξιώματα και αποδεικτικοί κανόνες επαρκούν για να απαντηθεί *οποιοδήποτε* μαθηματικό ερώτημα που μπορεί να εκφραστεί τυπικά σε αυτά τα συστήματα. Θα δειχθεί παρακάτω ότι αυτό δεν ισχύει … (Gödel [1931] 1967, σ. 596-597)

Εβδομήντα πέντε χρόνια μετά τη δημοσίευση του άρθρου, δεν έχει λήξει η συζήτηση για το ποιες είναι οι ευρύτερες συνέπειες των ΘΜΠ. Από μια άποψη, αυτό έπρεπε να αναμενόταν επειδή ακριβώς πρόκειται για *φιλοσοφική* συζήτηση. Και δεν εννοώ με αυτό ότι δεν υπάρχει πρόοδος στη φιλοσοφία. Πιστεύω ότι υπάρχει. Απλώς ένα μαθηματικό θεώρημα δεν μπορεί να *«λύσει οριστικά»* ένα φιλοσοφικό πρόβλημα. Ένα μαθηματικό θεώρημα περιορίζει απλώς τον χώρο των φιλοσοφικών δυνατοτήτων από την πλευρά της λογικο-μαθηματικής συνέπειας. Στηρίζει υποθετικά συμπεράσματα της μορφής: «*Αν* έχεις την *τάδε* φιλοσοφική θέση που εκφράζεται μαθηματικά με τον *τάδε* τρόπο, τότε πρέπει να έχεις και την *δείνα* φιλοσοφική θέση *εφόσον* αυτή εκφράζεται μαθηματικά με τον *δείνα* τρόπο». Όμως η αρχική ένταξη του προβλήματος στο πλαίσιο της μαθηματικής θεωρίας και, συνακόλουθα, η φιλοσοφική αποτίμηση του αποτελέσματος θα αποτελούν πάντοτε αντικείμενο ερμηνείας ή διαμάχης. Πάντως, όσο ανελεύθερο και αν ακούγεται, ο *περιορισμός των δυνατοτήτων συνιστά πρόοδο*.

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

Αναπολιτάνος, Δ. Α. (1985): *Εισαγωγή στη Φιλοσοφία των Μαθηματικών*. 4η έκδοση βελτιωμένη. Αθήνα: Εκδόσεις Νεφέλη.

Ρουσόπουλος, Γ. (1988): *Μελέτες για τον Εμπειρισμό*. Αθήνα: Ινστιτούτο του Βιβλίου – Α. Καρδαμίτσα.

Τζουβάρας, Αθ. (1992): *Στοιχεία Μαθηματικής Λογικής*. Θεσσαλονίκη: Εκδόσεις Π. Ζήτη.

Χριστοδουλίδης, Π. (επιμ.) (1993): *Η Φιλοσοφία των Μαθηματικών*. Αθήνα: Εκδόσεις Γ. Α. Πνευματικού.

Ayer, A. J. (1946): *Language, Truth and Logic*. 2nd edition. New York: Dover, 1952.

Azuma, H. (2006): “Two Deconstructions”. Αδημοσίευτο στα αγγλικά. Πηγή: Internet.

Benacerraf, P. (1967): “God, the Devil, and Gödel”, *The Monist* **51**: 9-32.

Benacerraf, P. and Putnam, H. (eds.) (1983): *Philosophy of Mathematics: Selected Readings*. 2nd edition. Cambridge: Cambridge University Press.

Blass, A. and Gurevich, Y. (2003): “Algorithms: A Quest for Absolute Definitions”, *Bulletin of European Association for Theoretical Computer Science* **81**:195-225.

Boolos, G. S. and Jeffrey, R. C. (1980): *Computability and Logic*. 2nd edition. Cambridge: Cambridge University Press.

Bouveresse, J. ([1999] 2002): *Γοητευτικές και Παραπλανητικές Ακροβασίες της Φιλοσοφίας*. Πρόλογος: Ν. Παναγιωτόπουλος. Μετάφραση: Τ. Μπούκη. Επιμέλεια: Σ. Βιρβιδάκης. Αθήνα: Εκδόσεις Πατάκη.

Carnap, R. ([1934] 1937): *The Logical Syntax of Language*. Translated from the German *Logische Syntax der Sprache* by A. Smeaton. London: Routledge & Kegan Paul.

Carnap, R. ([1950] 1988): “Εμπειρισμός, Σημασιολογία και Οντολογία” στο Γ. Ρουσόπουλος (1988), σ. 31-61. Πρώτη δημοσίευση με τίτλο “Empiricism, Semantics, and Ontology” στο *Revue Internationale de Philosophie* **4** (1950): 20-40. Ανατυπώθηκε με μικρές αλλαγές στο R. Carnap (1956), *Meaning and Necessity: A Study in Semantics and Modal Logic*. 2nd edition. Chicago: The University of Chicago Press, σ. 205-221. Αυτή η εκδοχή βρίσκεται και στο P. Benacerraf and H. Putnam (eds.) (1983), σ. 241-257 και χρησιμοποιήθηκε για απόδοση στα ελληνικά από τον Γ. Ρουσόπουλο.

Crocco, G. (2003): “Gödel, Carnap and the Fregean Heritage”, *Synthese* **137**: 21-41.

Detlefsen, M. (2005): “Formalism” στο S. Shapiro (ed.) (2005), σ. 236-317.

Enderton, H. B. (1972): *A Mathematical Introduction to Logic*. New York: Academic Press.

Friedman, M. (1999): *Reconsidering Logical Positivism*. Cambridge: Cambridge University Press.

Glymour, C. (1992): *Thinking Things Through: An Introduction to Philosophical Issues and Achievements*. Cambridge, MA: MIT Press.

Gödel, K. ([1930] 1967): “The completeness of the axioms of the functional calculus of logic” στο J. van Heijenoort (ed.) (1967), σ. 582-591.

Gödel, K. ([1931] 1967): “On formally undecidable propositions of *Principia Mathematica* and related systems I” στο J. van Heijenoort (ed.) (1967), σ. 592-616.

Gödel, K. ([1953/9] 1995): “Is Mathematics Syntax of Language? [Versions III, V]” στο S. Feferman, J. Dawson, W. Goldfarb, C. Parsons, and R. Solovay (eds.), *Kurt Gödel: Collected Works, volume III*. Oxford: Oxford University Press, σ. 334-362.

Gödel, K. ([1964] 1993): “Τι είναι το πρόβλημα του συνεχούς του Cantor;” στο Π. Χριστοδουλίδης (επιμ.) (1993), σ. 181-202. Απόδοση στα ελληνικά του “What is Cantor’s continuum problem?” που περιέχεται στο P. Benacerraf and H. Putnam (eds.) (1983), σ. 470-485.

Haugelend, J. ([1989] 1992): *Τεχνητή Νοημοσύνη. Σχεδιάζοντας τη νόηση: από την υπολογιστική θεωρία στις σύγχρονες ευφυείς μηχανές*. Μετάφραση: Σ. Ζαχαρίου. Επιστημονική επιμέλεια: Σ. Μανουσέλης. Αθήνα: Εκδόσεις Κάτοπτρο.

Hedman, S. (2006): *A First Course in Logic: An Introduction to Model Theory, Proof Theory, Computability, and Complexity*. Oxford: Oxford University Press.

Hilbert, D. ([1899] 1992): *Foundations of Geometry*. 2nd English edition. Translated from the 10th German edition of *Grundlagen der Geometrie* by L. Unger. Revised and enlarged by P. Bernays. La Salle, IL: Open Court.

Hilbert, D. ([1926] 1993): “Για το άπειρο” στο Π. Χριστοδουλίδης (επιμ.) (1993), σ. 139-163. Πρώτη δημοσίευση με τίτλο “Über das Unendliche” στο *Mathematische Annalen* **95** (1926): 161-190. Απόδοση στα αγγλικά από τον S. Bauer-Mengelberg στο J. van Heijenoort (ed.) (1967), σ. 367-392. Απόδοση στα αγγλικά από τους E. Putnam και G. J. Massey στο P. Benacerraf and H. Putnam (eds.) (1983), σ. 183-201.

Hilbert, D. ([1927] 1967): “The Foundations of Mathematics” στο J. van Heijenoort (ed.) (1967), σ. 464-479.

Hobbes, T. ([1651] 1985): *Leviathan*. Edited with an introduction by C. B. Macpherson. London: Penguin Books.

Kim, J. ([1998] 2005): *Η Φιλοσοφία του Νου*. Μετάφραση: Ε. Μανωλακάκη. Επιστημονική επιμέλεια: Σ. Ψύλλος. Αθήνα: Εκδόσεις Leader Books.

Lucas, J. (1961): “Minds, Machines and Gödel”, *Philosophy* **36**: 112-127.

McCall, S. (1999): “Can a Turing machine know that the Gödel sentence is true?” *Journal of Philosophy* **96**: 525-532.

McCall, S. (2001): “On ‘seeing’ the truth of Gödel sentence”, *Facta Philosophica* **3**: 25-29.

Moore, A. W. (1988): “What Does Gödel’s Second Incompleteness Theorem Show?” *Noûs* **22**: 573-584.

Moschovakis, Y. N. (2001): “What Is an Algorithm?” στο B. Engquist and W. Schmid (eds.), *Mathematics Unlimited – 2001 and Beyond*. New York: Springer, σ. 919-936.

Penrose, R. (1989): *The Emperor’s New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford: Oxford University Press.

Penrose, R. (1994): *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford: Oxford University Press.

Raatikainen, P. (2005): “On the Philosophical Relevance of Gödel’s Incompleteness Theorems”, *Revue Internationale de Philosophie* **234**: 513-534.

Rucker, R. ([1982] 1999): *Το Άπειρο και ο Νους*. Απόδοση στα ελληνικά – επιστημονική επιμέλεια: Κ. Χατζηκυριάκου. Ηράκλειο: Πανεπιστημιακές Εκδόσεις Κρήτης.

Shapiro, S. (2000): *Thinking about Mathematics*. Oxford: Oxford University Press. Έχει εκδοθεί με τίτλο *Σκέψεις για τα Μαθηματικά* (μετάφραση: Αθ. Δρόσος, Δ. Σπανός – επιστημονική επιμέλεια: Αθ. Δρόσος) από τις εκδόσεις του Πανεπιστημίου Πατρών, 2006.

Shapiro, S. (ed.) (2005): *The Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford: Oxford University Press.

Smullyan, R. (2001): “Gödel’s Incompleteness Theorems” στο L. Goble (ed.), The *Blackwell Guide to Philosophical Logic*. Oxford: Blackwell, σ. 72-89.

Tennant, N. (2001): “On Turing machines knowing their own Gödel sentences”, *Philosophia Mathematica* **9**: 72-79.

Van Heijenoort, J. (ed.) (1967): *From Frege to Gödel: A Sourcebook in Mathematical Logic, 1879-1931*. Cambridge, MA: Harvard University Press.

Zach, R. (2003): “Hilbert’s Program” στο E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/hilbert-program/.

Zach, R. (2006): “Hilbert’s Program Then and Now” στο D. Jacquette (ed.), *Philosophy of Logic*. Handbook of the Philosophy of Science, vol. 5. Amsterdam: Elsevier, σ. 411-447.

ΣΗΜΕΙΩΣΕΙΣ

[*] Διάλεξη για τα εκατό χρόνια από τη γέννηση του Kurt Gödel στο πλαίσιο εκδηλώσεων της εταιρείας μαθηματικών «Θαλής και Φίλοι», Νέο Κτήριο Μουσείου Μπενάκη, 21 Νοεμβρίου 2006. Ευχαριστώ θερμά τα μέλη της εταιρείας «Θαλής και Φίλοι» για την πρόσκληση και την τεχνική υποστήριξη.

[1] Πρόκειται για το Gödel ([1931] 1967).

[2] Πρόκειται για το Gödel ([1930] 1967) το οποίο συνοψίζει τη διδακτορική διατριβή του Gödel.

[3] Για μια πληρέστερη επισκόπηση βλ. Raatikainen (2005).

[4] Μια πρόταση που αποδεικνύεται τυπικά στο πλαίσιο μιας πρωτοβάθμιας θεωρίας είναι αληθής σε κάθε *μοντέλο* της θεωρίας αυτής αλλά μπορεί να είναι ψευδής σε μια *ερμηνεία* της γλώσσας της θεωρίας που *δεν είναι* μοντέλο της θεωρίας.

[5] Η παρουσίαση που ακολουθεί δεν διεκδικεί εύσημα μαθηματικής αυστηρότητας. Για παράδειγμα, συχνά χρησιμοποιώ απλώς τον όρο ‘θεωρία’ αντί του ‘πρωτοβάθμια θεωρία’, τον όρο ‘αναδρομική συνάρτηση’ αντί του ‘μερική αναδρομική συνάρτηση’, κ.λπ., και δεν δίνω ορισμούς κεντρικών όρων όπως ‘μηχανή Turing’, κ.ά. Ωστόσο, ελπίζω να είναι αρκετά αυστηρή ώστε να επιτρέψει τον εντοπισμό διαφόρων παρανοήσεων σχετικά με το περιεχόμενο των θεωρημάτων μη πληρότητας. Ο αναγνώστης που επιζητά μαθηματική αυστηρότητα ας συμβουλευτεί τα Τζουβάρας (1992), Boolos and Jeffrey (1980), Enderton (1972) ή Hedman (2006). Σύντομες και εύληπτες παρουσιάσεις των θεωρημάτων μη πληρότητας περιέχονται στα Rucker ([1982] 1999) και Smullyan (2001).

[6] Βλ., π.χ., Moschovakis (2001) και Blass and Gurevich (2003).

[7] Η περιγραφή αυτή δεν είναι παρά σύνθεση εκείνων που περιέχονται στο Glymour (1992, σ. 302) και Haugeland ([1989] 1992, σ. 93)

[8] Εξ ορισμού, η χαρακτηριστική συνάρτηση χ_A ενός υποσυνόλου A ενός συνόλου Ω έχει πεδίο ορισμού το Ω και απεικονίζει τα στοιχεία του Ω που δεν ανήκουν στο A στην τιμή 0 ενώ τα στοιχεία του A στην τιμή 1.

[9] Πρόκειται για τη θεωρία Q στο Boolos and Jeffrey (1980, σ. 158).

[10] Χρησιμοποιώντας αδόκιμο συμβολισμό, ας πούμε μόνο το εξής: μια θεωρία T που επεκτείνει την PA είναι ω-συνεπής αν και μόνο αν δεν υπάρχει τύπος $\varphi(x)$ τέτοιος ώστε οι $\varphi(0), \varphi(1), \varphi(2), \ldots$ να είναι αποδείξιμοι στην T και επίσης ο τύπος $\exists x \neg \varphi(x)$ να είναι αποδείξιμος στην T. Η ίδια η PA

είναι ω-συνεπής. Και αν μια θεωρία είναι ω-συνεπής, τότε θα είναι συνεπής. Όμως το αντίστροφο δεν ισχύει.

[11] Βλ., π.χ., Boolos and Jeffrey (1980, σ. 219-227).

[12] Για μια φιλοσοφική ανάλυση του 2ουΘΜΠ βλ. και Moore (1988).

[13] Κατά την κλασική λογική κάθε επιχείρημα με ασυνεπές σύνολο προκειμένων είναι έγκυρο ανεξάρτητα από το ποιο είναι το συμπέρασμα του επιχειρήματος.

[14] Η προκλητική αυτή αναλογία οφείλεται στον Smullyan (2001, σ. 84).

[15] Ο Rucker ([1982] 1999, σ. 303) ισχυρίζεται ότι ένα μυθιστόρημα που θα απαρτιζόταν από τη μοναδική πρόταση «Ο Γιάννης Χ δεν υπάρχει» θα αποτελούσε μια *πλήρη* περιγραφή του Γιάννη Χ.

[16] Πρόκειται για το αξίωμα το οποίο ο Hilbert ([1899] 1992, σ. 25) διατυπώνει ως εξής: «Έστω α μια ευθεία και Α ένα σημείο εκτός αυτής. Τότε υπάρχει το πολύ μια ευθεία στο επίπεδο, που προσδιορίζουν η α και το Α, η οποία περνά από το Α και δεν τέμνει την α».

[17] Για μια εισαγωγή στη πρόγραμμα του Hilbert, ο αναγνώστης μπορεί να ανατρέξει στα Αναπολιτάνος (1985, σ. 244-261) και Shapiro (2000, σ. 158-168). Οι πιο απαιτητικοί ας συμβουλευτούν τα Zach (2003, 2006) και, για μια ευρύτερη σκοπιά, το Detlefsen (2005). Το κείμενο της διάλεξης που έδωσε ο Hilbert ([1926] 1993) προς τιμήν της μνήμης του Weierstrass στη Μαθηματική Εταιρεία της Βεστφαλίας στις 4 Ιουνίου του 1925 και που δημοσιεύτηκε την επόμενη χρονιά στο *Mathematische Annalen* με τίτλο "Über das Unendliche" είναι η πληρέστερη παρουσίαση των βασικών ιδεών του προγράμματος από τον ίδιο τον εμπνευστή του. Αποδόσεις αυτού του κειμένου στα αγγλικά βρίσκονται στις συλλογές Benacerraf and Putnam (eds.) ([1964] 1983, σ. 183-201) και van Heijenoort (1967, σ. 367-392).

[18] Οφείλω αυτή την παράφραση του Dickens στον Shapiro (2000, σ. 158).

[19] Ανάλογη πρόκληση αντιμετωπίζει κάθε φορμαλιστική φιλοσοφία των μαθηματικών που «διαιρεί» τα μαθηματικά σε μια περιοχή στην οποία ορισμένοι τύποι πρέπει να αντιμετωπίζονται ως προτάσεις με νόημα και σε μια περιοχή στην οποία κάθε τύπος πρέπει να αντιμετωπίζεται εργαλειοκρατικά ως αλγεβρικο-συμβολικό μόρφωμα. Ο Detlefsen (2005, σ. 306-309) αποδίδει το όνομα «πρόβλημα διαίρεσης» ["division problem"] σε αυτήν την πρόκληση.

[20] Βλ. και Hilbert ([1927] 1967, σ. 469-471).

[21] Ορισμένοι έχουν ισχυριστεί ότι και το 1οΘΜΠ του Gödel υψώνει ανυπέρβλητα εμπόδια για το πρόγραμμα του Hilbert. Αλλά για λόγους οικονομίας που επιβάλλει ο χαρακτήρας του παρόντος κειμένου δεν θα αναφερθώ καθόλου σε αυτή την πτυχή της θεματικής. Ο ενδιαφερόμενος αναγνώστης μπορεί να ανατρέξει στις βιβλιογραφικές πηγές που παρατίθενται στη σημείωση 11.

[22] Κυρίως από τον Detlefsen. Βλ. Detlefsen (2005), Zach (2003, 2006) και τις βιβλιογραφικές αναφορές που παρατίθενται εκεί.

[23] Χρησιμοποιώ τους δυο όρους ως συνώνυμους.

[24] Βλ. και Ayer ([1946] 1952, κεφ. IV).

[25] Ας σημειωθεί ότι η αδρομερής αυτή παρουσίαση δεν λαμβάνει υπόψη ούτε τη μεταβολή ορισμένων απόψεων του Carnap για τα μαθηματικά ούτε την αυστηρότητα ορισμένων διατυπώσεων του Gödel. Για τέτοιες «λεπτομέρειες», ο ενδιαφερόμενος αναγνώστης μπορεί να συμβουλευτεί το άρθρο της Crocco (2003). Συναφή είναι και τα άρθρα που περιλαμβάνονται στο τρίτο μέρος του βιβλίου του Friedman (1999).

[26] Για συντομία, χρησιμοποιώ εδώ και στο εξής τον όρο 'γκεντελιανό σύστημα' για κάθε αξιωματικοποιήσιμη πρωτοβάθμια θεωρία που επεκτείνει την PA. Η καθιερωμένη διάκριση ανάμεσα σε «αξιώματα» και «κανόνες» μπορεί εδώ να αγνοηθεί δεδομένης της διαδεδομένης, ανάμεσα σε λογικούς θετικιστές, άποψης ότι τα «αξιώματα», ακόμη και εκείνα *φυσικών* θεωριών, δεν είναι παρά «εισιτήρια συναγωγών» - «άδειες» που χορηγεί η θεωρία για τη συναγωγή ορισμένων προτάσεων από άλλες προτάσεις.

[27] Βλ., π.χ., Glymour (1992, κεφ. 13) και Kim ([1998] 2005, κεφ. 4).

[28] Στο παράθεμα, 'Maud' είναι απλώς το όνομα κάποιας μηχανής. Έχω αλλάξει μερικούς συμβολισμούς και συντομογραφίες του πρωτοτύπου ώστε να συμφωνούν με αυτούς και αυτές που έχω υιοθετήσει εδώ. Τέλος η έκφραση «συνήγοροι των δυνατοτήτων» των μηχανών είναι μάλλον παραπλανητική. Η κριτική των επιχειρημάτων των Lucas, Penrose ή McCall δεν αποσκοπεί στην τεκμηρίωση της θέσης ότι η ανθρώπινη νόηση μπορεί να προσομοιωθεί στη λειτουργία μηχανών αλλά στην υπονόμευση του ισχυρισμού ότι τα ΘΜΠ, ενταγμένα στα εν λόγω επιχειρήματα, καταδεικνύουν ότι *δεν* μπορεί να προσομοιωθεί.

[29] Από το Davies, P. (1992): *The Mind of God*. New York: Simon & Schuster. Παρατίθεται στο Raatikainen (2005).

[30] Η εικασία του Cantor λέει ότι δεν υπάρχει (άπειρο) σύνολο με πληθικό αριθμό αυστηρά μεταξύ του πληθικού αριθμού του συνόλου των φυσικών αριθμών και εκείνου του συνόλου των πραγματικών αριθμών (δηλαδή, του συνεχούς).

[31] Ευχαριστώ τον Στέλιο Βιρβιδάκη που μου συνέστησε το βιβλίο αυτό.

[32] Βλ. Bouveresse ([1999] 2002, σ. 119).