

Applicare le “leggi” di Asimov

Giancarlo Livraghi – marzo 2007

Devo confessare che non ci avevo mai pensato. E non mi sarebbe venuto in mente se non avessi visto una serie di vignette satiriche dello stesso autore che ho citato in *La stupidità delle tecnologie*¹ – e anche in varie altre occasioni precedenti.

Alla fine di questo articolo (pagine 6-10) si trovano, con un breve commento, le cinque vignette pubblicate da J.D. Frazer (“Illiad”) dal 19 al 23 marzo 2007 più una che ha aggiunto il 3 aprile e altre cinque dal 17 al 21 aprile.

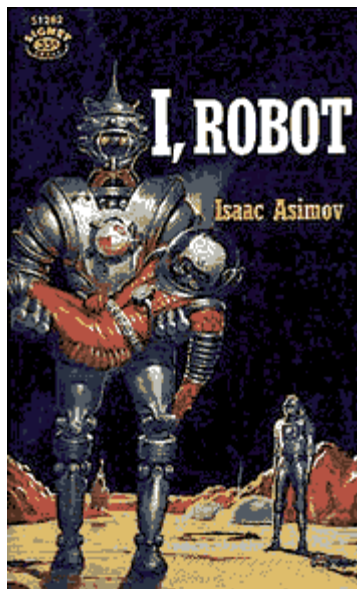
Si tratta di capire se possiamo applicare alle tecnologie di oggi le *Three Laws of Robotics* concepite da Isaac Asimov e poi seguite non solo da lui, ma anche da molti altri autori, come norme fondamentali di comportamento per macchine che si immaginano programmate in modo da avere qualche capacità di “pensare”.²

Da più di cinquant’anni quelle “leggi” sono generalmente accettate non solo nel mondo della *science fiction*, ma anche negli studi scientifici di cibernetica e di ipotetica “intelligenza artificiale”.

Furono definite da Asimov in un racconto, *Runaround*, del 1942, e poi in una serie di sviluppi sul tema dei robot, immaginati non come “androidi” (cioè macchine cibernetiche di aspetto e comportamento umano) ma come “esseri” palesemente meccanici di forma approssimativamente antropomorfa – come vediamo, per esempio, in questa immagine, che è la copertina della prima edizione di *I, Robot* in cui era raccolta la prima serie di racconti di Asimov su questo tema, dal 1940 al 1950.

¹ <http://gandalf.it/offline/tecstup.htm> (anche pdf)

² Vedi le spiegazioni in http://it.wikipedia.org/wiki/Tre_leggi_della_robotica



Quest'altra immagine è molto più recente.
È un prototipo di robot che si chiama "Asimo"
in onore di Isaac Asimov.



Lo stile è diverso. Ma il concetto è lo stesso.

Naturalmente i robot esistono, ma non si è sviluppata quell'ipotesi su cui si basavano le narrazioni del "possibile": un'estesa proliferazione di macchine con questo aspetto, più o meno "pensanti", al servizio dell'umanità – capaci di alleggerire il genere umano dei compiti più gravosi, o pericolosi, o meno piacevoli, o comunque "servili".

È ovvio che, se quelle macchine esistessero, dovrebbero esserci criteri “universalmente” validi, e chiaramente definiti, per regolare il loro comportamento. Che si identificano con le “tre leggi della robotica”. Queste norme inderogabili sono in ordine rigorosamente “gerarchico”: cioè la prima prevale sulle altre due e la seconda sulla terza.

1. Un robot non deve mai recare danno a un essere umano né, per inattività, permettere che un essere umano subisca un danno.

Usciamo da un mondo immaginario popolato di miliardi di robot più o meno antropomorfi – e vediamo che cosa succede nella nostra realtà quotidiana.

Le macchine (e in generale gli automatismi, anche quando ad applicarli sono esseri umani) obbediscono alla “prima legge”? No. Non solo i computer, ma anche aggeggi più semplici, non sono “buoni servitori”. Ci infliggono continuamente danni di ogni sorta. E poco o nulla fanno per impedire che qualcosa o qualcun altro ci danneggi. La gamma delle conseguenze varia da piccoli fastidi quotidiani a danni estremamente gravi.

Questi meccanismi sono essenzialmente stupidi. Se si limitassero a svolgere fedelmente compiti semplici e coerenti, potrebbero esserci utili e darci poco fastidio. Il problema è una tendenza delirante a renderli sempre più complessi, sempre più carichi di funzioni inutili o inutilmente pasticciate.

Lo sviluppo e l'applicazione di ogni sorta di tecnologie potrebbero avere un sostanziale miglioramento se tutta la progettazione fosse severamente assoggettata alla Prima Legge della Robotica.

2. Un robot deve obbedire agli ordini ricevuti da un essere umano, se tali ordini non contravvengono alla Prima Legge.

Le tecnologie (come gli automatismi burocratici, anche quando sono applicati da esseri umani) obbediscono sempre meno agli ordini (e alle esigenze) delle persone di cui dovrebbero essere al servizio. Sono dominate dalle intenzioni di progettisti che badano troppo spesso alle proprie fantasie tecnologiche (o privilegi e pregiudizi di categoria) e troppo raramente alle esigenze reali degli utilizzatori.

Inoltre gli stessi progettisti spesso non si rendono conto di quello che stanno facendo, perché nell'incrocio di sistemi sempre più complessi si trovano a usare routine di cui non conoscono la struttura – o a intersecare funzioni progettate da altri per motivi diversi. L'intrico che ne deriva è una complicazione indecifrabile, con una moltiplicazione di disfunzioni di cui spesso anche i tecnici più esperti non capiscono la causa (e perciò non trovano la soluzione).

Il risultato è che le tecnologie obbediscono sempre meno agli esseri umani – e si pretende sempre di più che gli esseri umani obbediscano alle tecnologie. Prima che gran parte dell'umanità abbia bisogno di cure psichiatriche, e che si producano disastri più palesemente catastrofici di quelli già avvenuti, sarebbe opportuno sottoporre ogni sviluppo tecnico (e disciplina burocratica) a una rigorosa e intransigente applicazione della Seconda Legge della Robotica.

Molte delle tecnologie oggi esistenti sarebbero da buttar via, per sostituirlle con cose molto più semplici e funzionali. (Magari ripescandone qualcuna che c'era, ma è stata sostituita da soluzioni meno valide).

Forse sarebbe un po' a disagio chi oggi guadagna, o esercita potere, infliggendoci quei pastrugni inestricabili che ogni giorno ci rovinano la vita e il buonumore. Ma starebbe molto meglio tutto il resto dell'umanità.

3. Un robot deve proteggere la propria esistenza, a condizione che tale protezione non sia in conflitto con la Prima e la Seconda Legge.

Nella realtà in cui viviamo... le tecnologie e le macchine sanno almeno proteggere se stesse? No. Nel percorso demenziale della complicazione sono sempre più soggette a guasti, malfunzionamenti, crisi e conflitti interni (nel senso che ogni funzione aggiunta rischia di guastarne un'altra).

Nella continua ricerca di falsa "innovazione" è sempre meno possibile aggiustarle. Quando qualcosa si guasta, ci sentiamo dire sempre più spesso che l'unica soluzione è buttarla via e sostituirla con una "nuova" (che, essendo più complicata, funziona peggio e si guasta di più). Quindi le macchine e le tecnologie, oltre a farci danni e a non obbedirci, sono sempre più avviate a un suicidio collettivo. Già oggi i rottami sono un preoccupante inquinamento ambientale.

Continuare su questo percorso non è intelligenza – né naturale, né artificiale. È stupidità, con componenti masochistiche. Prima che sia troppo tardi, rileggiamo le Leggi di Asimov e cerchiamo di applicarle.

In uno degli sviluppi successivi ai suoi primi lavori su questo tema, Isaac Asimov si era accorto che mancava un principio più generale, e perciò aveva aggiunto una “legge zero”, gerarchicamente superiore alle altre tre: «Un robot non deve danneggiare l’umanità né, per inattività, permettere che l’umanità subisca un danno».

Bene, facciamo così. Mettiamo le esigenze umane al di sopra di ogni altra considerazione – e vediamo come ridurre finalmente all’obbedienza le macchine con cui sarebbe molto più piacevole convivere se fossero davvero al nostro servizio.

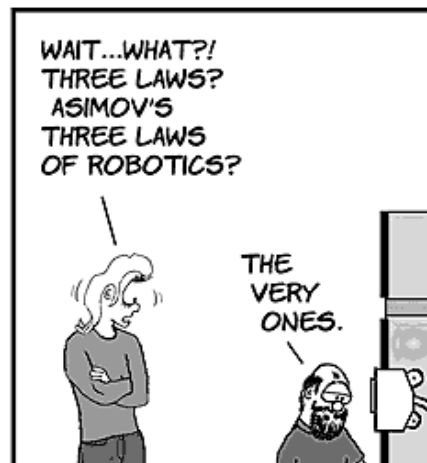
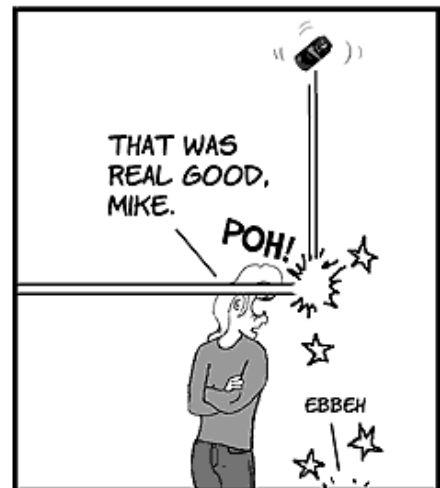
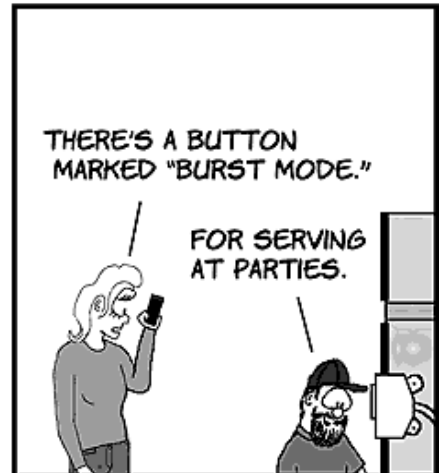
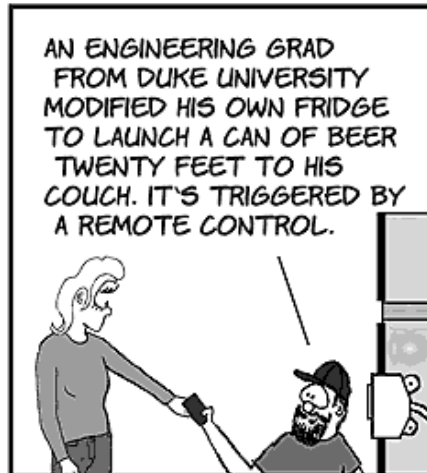
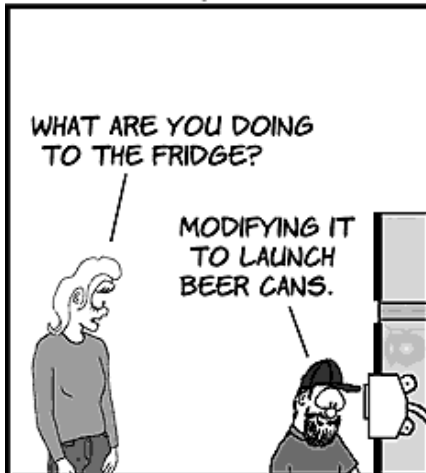
Post Scriptum

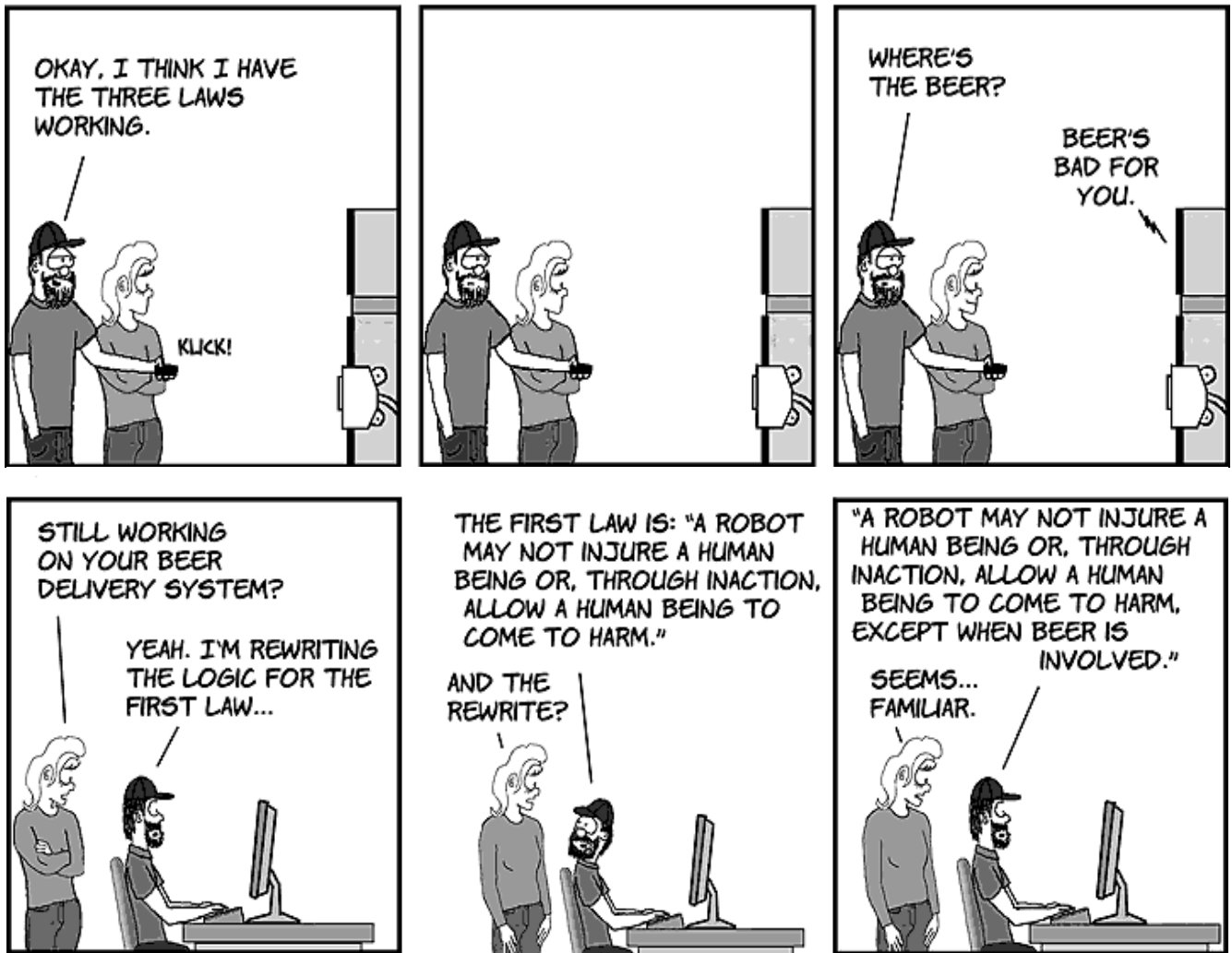
Sembra che qualcuno ci abbia già pensato
(ma solo per quanto riguarda i robot).

Secondo una notizia diffusa nel marzo 2007 il governo sudcoreano
intende stabilire una norma che impone le Leggi di Asimov
a tutti i robot esistenti o in corso di progettazione.

Appendice

Queste sono le cinque vignette pubblicate da J.D. Frazer
("Iliad" – <http://userfriendly.org>) dal 19 al 23 marzo 2007



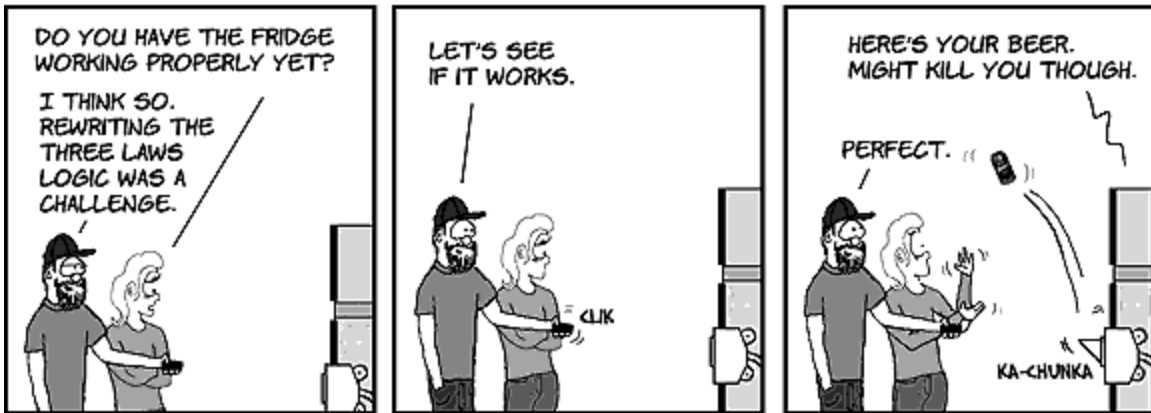


Copyright © J.D. Frazer "Illiad" – 2007

Miranda Corneille, che è il principale personaggio femminile in questo ambiente, chiede a Mike Floyd, che è un tecnico, che cosa sta facendo. Mike le spiega che ha deciso di programmare un frigorifero in modo da poter ricevere le lattine di birra senza alzarsi dal divano (il telecomando ha una variante *burst mode* per una distribuzione estesa quando c'è una festa).

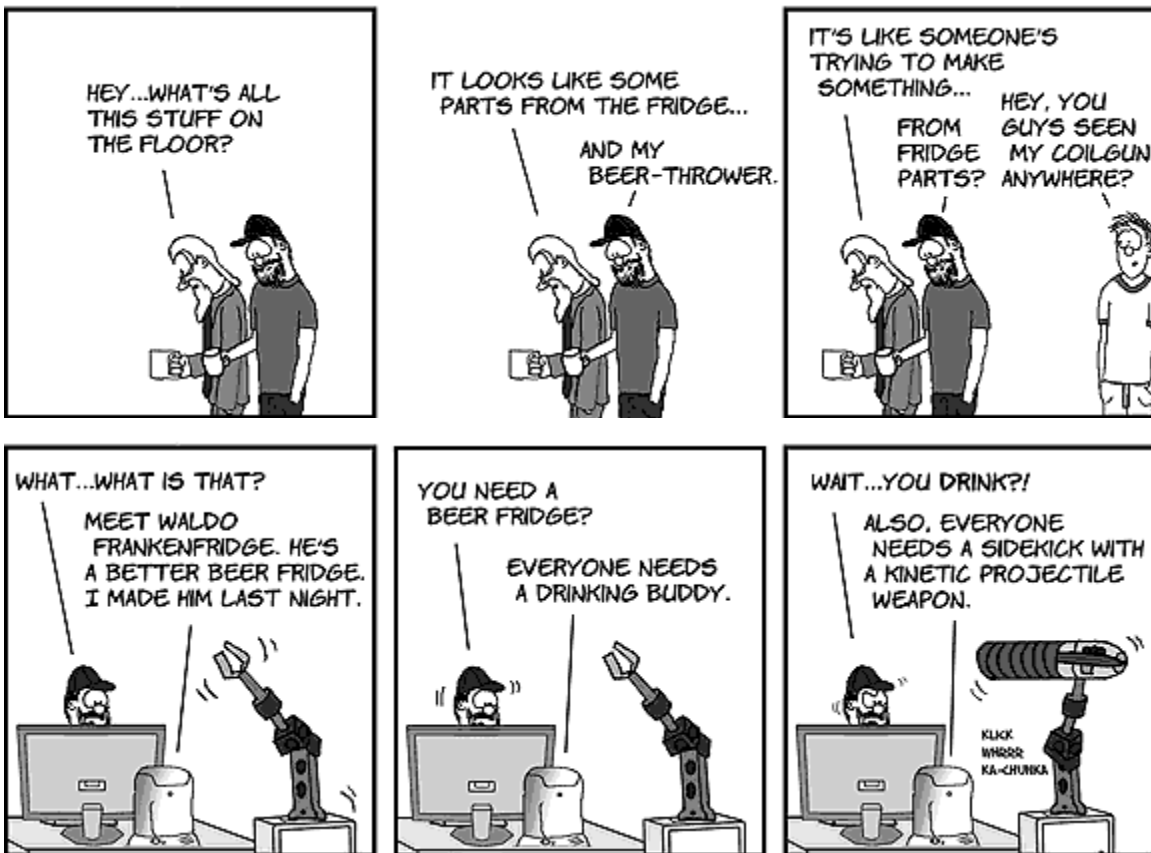
Constatato che il lancio delle lattine è pericoloso, Mike decide di inserire nel software la Prima Legge della Robotica. La conseguenza è che smette di funzionare "perché la birra può far male agli umani". Mike modifica la regola introducendo un'eccezione "fuorché quando c'entra la birra". Qui la serie di vignette finisce, ma è ovvio che in quel modo si ritorna al punto di partenza. La "morale della favola" è evidente.

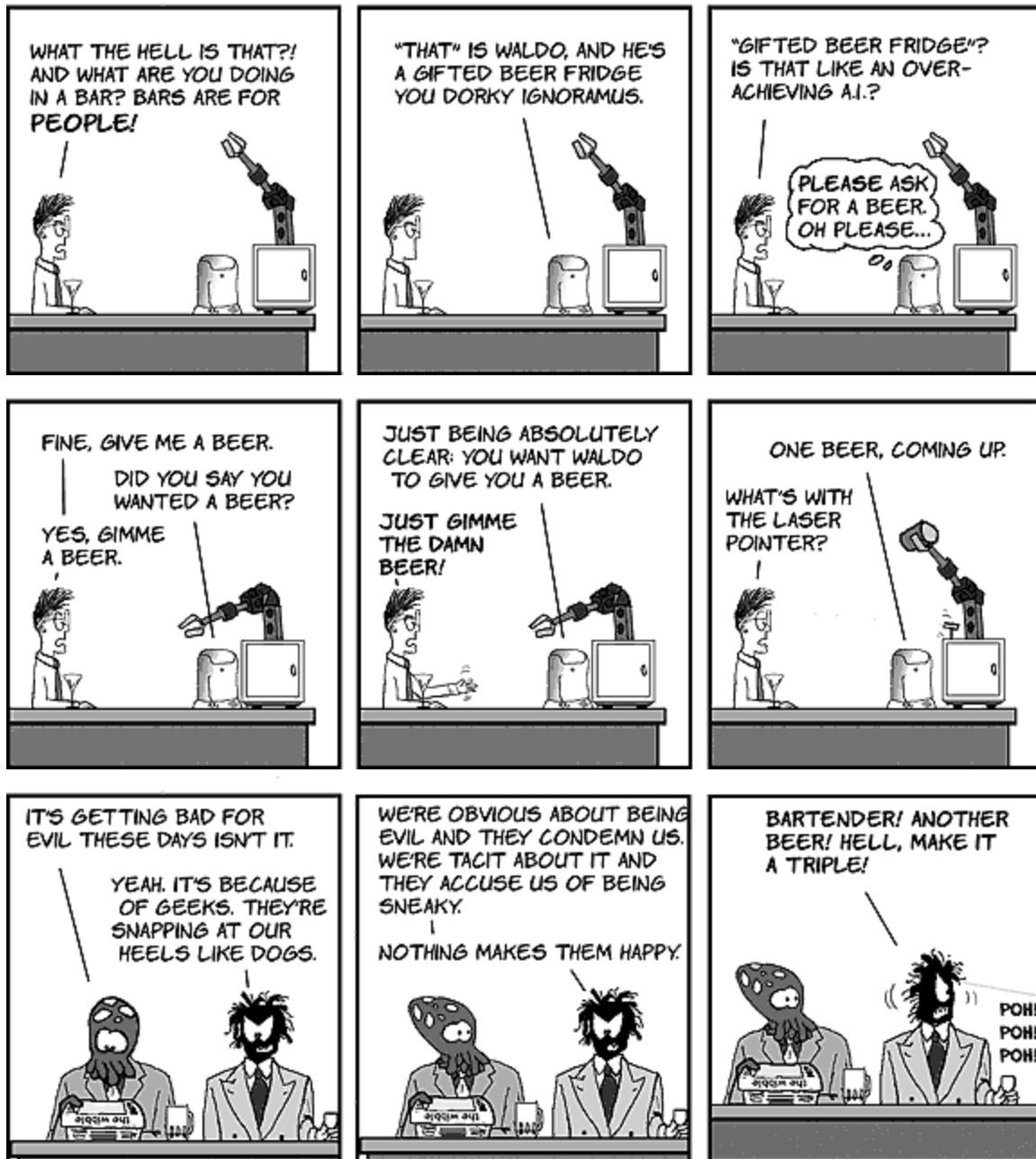
Ma più tardi, il 3 aprile 2007, "Illiad" ritorna sull'argomento.



Dopo vari tentativi Mike è riuscito a manipolare le Leggi di Asimov nel software che controlla il frigorifero. Così la macchina lancia la lattina, in modo meno traumatico, mentre emette un avviso: «ecco la tua la birra, ma ti potrebbe uccidere». Un’ovvia metafora non solo degli “avvisi” che si trovano sui pacchetti di sigarette e su ogni sorta di prodotti, ma anche dei compromessi con cui si possono eludere o “aggirare” le leggi.

Ma la piccola saga non è finita. Una nuova serie di cinque vignette sullo stesso tema si è svolta dal 17 al 21 aprile 2007.





Copyright © J.D. Frazer "Illiad" – 2007

Insieme a un altro personaggio, Sid Dabster, Mike trova in giro frammenti dispersi del frigorifero lancialattine, mentre A.J. Garrett si chiede perché sia scomparso il suo cannoncino a binario magnetico. Si scopre che è entrato in gioco un computer (si chiama Erwin ed è un'immaginaria forma spontanea di "intelligenza artificiale").

Erwin ha deciso di “migliorare” l’invenzione – e ha costruito una macchina con il nome preoccupante di Waldo Frankenfridge (frigo Frankenstein).

E poi... perché Erwin è in un bar ad aspettare che il suo amico Stef Murky voglia una birra? Quando Stef la chiede, che cosa si mette in moto? Perché c’è un tracciante laser e dove sta puntando?

A questo punto la storia si complica. Nel bar ci sono due mostruosi personaggi, Cthulhu e Hastur (anche questi ricorrenti nelle vignette di “Iliad” – derivati dalle storie *horror* di Howard Phillips Lovecraft) che sono ironiche incarnazioni del “male” e si preoccupano dell’ambiguità del loro ruolo. ³ Hastur ordina una tripla birra ... [*qui la storia si interrompe*]

Sembra che l’autore abbia deciso
di non spiegare la situazione
e di lasciarla interpretare ai lettori.

Un’ipotetica “intelligenza artificiale”
ha dato alla macchina
un comportamento disumano?

O sono in gioco intenzionali
“malvagità” tecnologiche?

Se dalla fantasia ritorniamo alla realtà,
non è difficile constatare che spesso
i due fattori si mescolano, provocando
ogni sorta di complicazioni e di problemi.

³ Si lamentano di non avere vita abbastanza facile. Dicono: «se riveliamo quello che siamo ci condannano, se lo nascondiamo ci accusano di imbrogliare».