

Olimpiada Informatyczna — ocenianie rozwiązań

Przed organizatorami różnych konkursów informatycznych zawsze staje problem, jak oceniać prace ich uczestników. Idealnie byłoby, gdyby istniał sprawiedliwy, szybki, klarowny, oceniający prawdziwą wiedzę uczestników, a minimalizujący rolę przypadku, system oceny. Niestety, przynajmniej kilku z powyższych warunków nie można ze sobą pogodzić. Wszelkie próby „ręcznej” oceny rozwiązań, np. przez jakąś komisję, oprócz wątpliwości co do klarowności takiego sposobu oceny, skazane są z góry na niepowodzenie przy większych konkursach, z powodu zbyt długiej czasochłonności takiego procesu. Przykładowo w I etapie VIII Olimpiady Informatycznej wzięło udział 1534 osób, co przy 4 zadaniach, daje w kilka tysięcy rozwiązań do sprawdzenia. Innym sposobem jest wprowadzenie automatycznego sprawdzania, które jest całkowicie bezduszne wobec nawet minimalnych ludzkich pomyłek, jednak w pełni spełnia wszystkie pozostałe postulaty idealnego sprawdzania prac. Podczas Olimpiady wszystkie prace sprawdzane są automatycznie z wykorzystaniem specjalnie do tego napisanego oprogramowania. Jak już zostało wcześniej wspomniane może nie jest to sposób pozbawiony wad, lecz nic lepszego nie zostało do tej pory wymyślone.

Ocenianie

Na czym więc polega dokładnie owe maszynowe ocenianie rozwiązań? Większość zadań, które pojawiają się na Olimpiadzie, polega na wczytaniu danych, wykonaniu obliczeń i zapisania wyniku. W takim wypadku przygotowuje się zestaw kilku–kilkunastu specjalnie przygotowanych (czytaj: najbardziej złośliwych) danych testowych, na których kolejno uruchamiany jest program zawodnika. Określenie „uruchamiany” jest pewnym uproszczeniem, gdyż nad tym procesem czuwa pewna tajemnicza maszyna, której celem jest zagwarantowanie uczciwego i sprawiedliwego procesu oceniania. Każdy z przypadków testowych ma przypisany pewien limit czasowy oraz liczbę punktów, które można otrzymać za zaliczenie testu. Jeśli program zawodnika przekroczy ustalony limit czasowy, jego działanie jest przerywane. Podczas wykonania programu zawodnika mogą zdarzyć się następujące sytuacje:

- program dał poprawną odpowiedź i zakończył się w wyznaczonym limicie czasowym,
- program przekroczył dopuszczalny czas wykonania,
- nastąpił błąd wykonania,
- program udzielił błędnej odpowiedzi (lub jej nie udzielił).

Tylko w pierwszym przypadku (tzn. udzielenie prawidłowej odpowiedzi w wyznaczonym czasie) test jest zaliczany, we wszystkich pozostałych przypadkach rozwiązanie otrzymuje 0 punktów za test. Dokładna liczba punktów za test zależy od czasu wykonania. Jeśli maksymalny dopuszczalny czas wynosił T , a program zakończył się w czasie krótszym niż $\frac{T}{2}$, to otrzymuje maksymalną możliwą liczbę punktów; w przedziale $\frac{T}{2} \dots T$ uzyskane punkty liniowo maleją, tak by osiągnąć 0 przy czasie T . Zależność ta pokazana jest na rys. 1.

Rys. 1. Wykres przedstawiający zależność między otrzymanymi punktami a czasem wykonania.

Grupowanie testów

Czasami zachodzi potrzeba umieszczenia w testach danych, dla których odpowiedź jest bardzo łatwa do obliczenia, czy wręcz zgadnięcia (np. odpowiedź NIE). W takich wypadkach stosuje się grupowanie testów. Grupowanie polega na uzależnieniu oceny za pewien zbiór testów od poprawnego zaliczenia wszystkich testów z grupy. Dokładniej wygląda to tak, że zawodnik nie otrzymuje oceny za poszczególne testy z grupy, lecz za całą grupę. Ocenę dla grupy jest równa minimalnej liczbie uzyskanych punktów z testów wchodzących w jej skład.

Limity czasowe

Często pada pytanie jak duże są limity czasowe dla danego zadania. Niestety w przypadku Olimpiady Informatycznej (jak i większości innych konkursów tego typu) nie jest on z góry podawany. Ustalenie, czy program jest już wystarczająco efektywny, czy nie, jest częścią rozwiązania zadania. Można jednak powiedzieć, że w 90% przypadków dopuszczalny limit czasowy dla maksymalnych danych jest rzędu kilku sekund (jakkolwiek **nie jest** to regułą).

Celem limitów czasowych, jest umożliwienie odróżnienia rozwiązań o różnych złożonościach czasowych. Stąd rozwiązanie o złożoności oczekiwanej przez jurorów, lecz bez specjalnych optymalizacji, powinno pomyślnie przejść wszystkie testy. Jednak już rozwiązanie o gorszej złożoności, nawet ze znacznymi optymalizacjami (np. przepisaniem części kodu w assemblerze) powinno przekroczyć limity przynajmniej dla części testów.

Rozwiązania

W chwili obecnej (październik 2002) podczas I etapu Olimpiady ocenie podlegają kody źródłowe programów. Zgłaszane programy są kompilowane w sposób podany w *Zasadach organizacji zawodów* (podane są tam wersje używanych kompilatorów i dokładne polecenia używane do kompilowania rozwiązań). Jeśli rozwiązanie nie uda się skompilować w powyższy sposób jest ono automatycznie odrzucane. Ważną zmianą w stosunku do lat poprzednich jest zmiana sposobu wczytywania danych wejściowych i wyjściowych. Obecnie rozwiązania powinny wczytywać dane ze standardowego wejścia i zapisywać wynik na standardowe wyjście.

Co jest ważne, a co mniej

Rozwiązania sprawdzane są automatycznie, ważne jest więc, by rozwiązania nie utrudniały tego procesu.

Co jest ważne:

- Poprawny format odpowiedzi — rozwiązania powinny dokładnie stosować się do formatu wyjścia podanego w zadaniu, wszelkie niezgodności mogą zostać potraktowane jako nieprawidłowa odpowiedź.
- Sposób rozwiązania — rozwiązania niezbyt efektywne mogą otrzymać bardzo mało punktów. Cały sekret tkwi w tym, aby znaleźć algorytm, który potrafi rozwiązać zadany problem (nawet dla najbardziej złośliwych danych) w krótkim czasie. Tylko takie rozwiązanie ma szansę na uzyskanie 100% punktów.

A co trochę mniej:

- Styl programowania, komentarze — niestety nie da się tych rzeczy formalnie zdefiniować, ani obiektywnie ocenić, stąd nie mają one wpływu na ocenę rozwiązania. Jednak statystycznie rzecz biorąc rozwiązania niechlujnie zakodowane i pozbawione komentarzy zawierają więcej błędów niż takie, przy których dołożono przynajmniej pewnej staranności.
- Sposób rozwiązania — z powodów dokładnie takich jak w poprzednim punkcie, tak długo, jak rozwiązania zawodnika jest nie mniej efektywne niż rozwiązanie wzorcowe, jest tak samo dobre. Stąd tyle samo punktów może otrzymać rozwiązanie, które wypisuje jedynie wcześniej stabilizowane wartości, jak i takie, które szybko je oblicza. Jeśli zadanie jest tak sformułowane, że możliwe jest stabilizowanie odpowiedzi, to znaczy, że organizatorzy dopuścili taką formę rozwiązania.

Powodzenia!