Ngôn ngữ lập trình C

Bài 9. Cấu trúc và nhóm

Soạn bởi: TS. Nguyễn Bá Ngọc

Nội dung

- Khái niệm & cú pháp
- Các toán tử
- Xử lý cấu trúc và nhóm

Nội dung

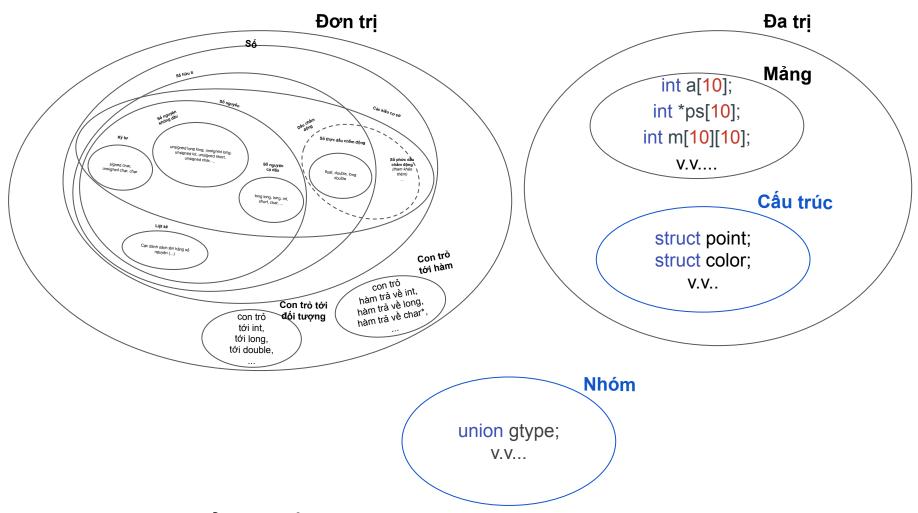
- Khái niệm & cú pháp
- Các toán tử
- Xử lý cấu trúc và nhóm

Các khái niệm

- Kiểu cấu trúc mô tả 1 danh sách không rỗng các đối tượng thành viên được lưu trong các phân vùng nhớ tuần tự theo thứ tự khai báo (nhưng không bắt buộc phải liên tục như mảng), trong vùng nhớ của đối tượng cấu trúc. Các thành viên có thể có kiểu khác nhau.
- Kiểu nhóm mô tả 1 tập không rỗng các đối tượng thành viên được lưu ở cùng 1 địa chỉ, đó cũng là địa chỉ của đối tượng nhóm. Các thành viên có thể có kiểu khác nhau.

Cấu trúc rỗng và nhóm rỗng (không chứa thành viên nào) không hợp lệ trong quy chuẩn ISO, nhưng hợp lệ trong mở rộng GNU (GCC).

Một số (lớp) kiểu dữ liệu trong C99



*Tuy có thể có nhiều thành viên nhưng nhóm không thuộc lớp đa trị bởi vì các thành viên chồng lấn bộ nhớ - một cách tương đối ở 1 thời điểm 1 đối tượng nhóm chỉ lưu được 1 giá trị của 1 thành viên.

Sơ lược cú pháp định nghĩa kiểu

- Cấu trúc: struct định-danh { /* Danh sách thành viên */ };
- Nhóm: union định-danh { /* Danh sách thành viên */ };
- Ví dụ:

```
// Thành viên y được cấp phát sau thành viên x chồng lấn bộ nhớ struct xy { union xy { long x; double y; }; // Kiểu cấu trúc rỗng struct empty {}; // Kiểu nhóm rỗng union empty {};
```

struct empty và union empty trong ví dụ không hợp lệ trong quy chuẩn ISO, nhưng hợp lệ trong mở rộng GNU (GCC).

Sơ lược cú pháp tạo đối tượng

- Khai báo biến cấu trúc:
 - Định dạng: struct tên-cấu-trúc tên-biến;
 - struct point p, q; // p và q là các (biến) cấu trúc, có kiểu struct point
 - const struct point cp, cq; // cp và cq là các cấu trúc chỉ đọc
- Khai báo biến nhóm (tương tự cấu trúc):
 - Định dạng: union tên-nhóm tên-biến;
 - union gtype a, b; // a và b là các (biến) nhóm, có kiểu union gtype
 - const gtype ca, cb; // ca và cb là các nhóm chỉ đọc
- Cũng có thể khai báo các biến khi định nghĩa kiểu:
 - struct point {double x, y; } p, q, *pp = &p;
 - pp là con trỏ và được khởi tạo trỏ tới p.
 - union gtype { long I; double d; } a, b, *pg = &a;
 - pg là con trỏ và được khởi tạo trỏ tới a.

!Lưu ý: Các từ khóa struct và union được kết hợp với định danh để xác định kiểu cấu trúc và kiểu nhóm.

Sơ lược cú pháp tạo con trỏ

Có thể áp dụng cơ chế suy diễn kiểu con trỏ tương tự như đã thực hiện với các kiểu đơn tri.

- struct point *pp;
 - o pp là con trỏ tới cấu trúc point, có kiểu struct point *
- const struct point *pcp;
 - con trỏ tới cấu trúc point chỉ đọc, có kiểu const struct point *
- struct point * const cpp; // bién chỉ đọc kiểu struct point *
 - con trỏ chỉ đọc tới cấu trúc point, có kiểu struct point *const
- union gtype *pg;
 - con trỏ tới nhóm gtype, pg có kiểu union xy *
- const union gtype *pcg;
 - con trở tới nhóm gtype chỉ đọc, có kiểu const union gtype*
- union gtype * const cpg; // biến chỉ đọc kiểu union gtype *
 - con trỏ chỉ đọc tới nhóm gtype, có kiểu union gtype * const ,

Sơ lược cú pháp tạo con trỏ₍₂₎

- Có thể tạo con trỏ tới cấu trúc và con trỏ tới nhóm trong ngữ cảnh không có các định nghĩa kiểu cấu trúc và kiểu nhóm tương ứng.
- Cấu trúc và nhóm có thể có thành viên là con trỏ tới chính nó, nhưng không thể có thành viên là chính nó:
 - struct node { long data; struct node *next; }; // Ok
 - struct node {long data; struct node nd; }; // Lỗi Kiểu struct node chưa hoàn thiện ở thời điểm khai báo nd.
 - union gtype { long l; double d; union gtype *pg; }; // Ok

Khởi tạo cấu trúc và nhóm

Có thể khởi tạo cấu trúc và nhóm bằng danh sách khởi tạo hoặc 1 đối tượng khác

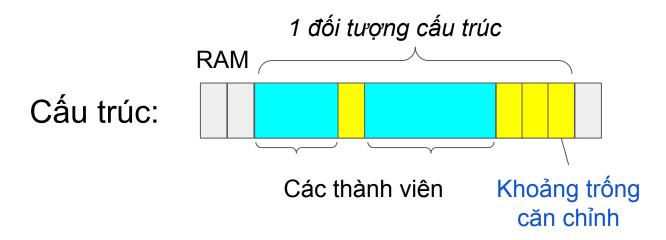
- Danh sách khởi tạo:
 - struct point p = {20, 30}; // Khởi tạo tuần tự
 - struct point p = {.x = 100}; // Khởi tạo thành viên x
 - Nếu cấu trúc được khởi tạo 1 phần thì các thành viên còn lại được khởi tạo với giá trị mặc định tương tự biến toàn cục (= 0 theo kiểu của thành viên).
 - o union gtype g = {.i = 101}; // Khởi tạo thành viên i của g
- Khởi tạo bằng đối tượng khác:
 - struct point p1 = p; // Khởi tạo p1 giống như p
 - union gtype g1 = g; // Khởi tạo g1 giống như g

Các giá trị gộp và đối tượng không tên

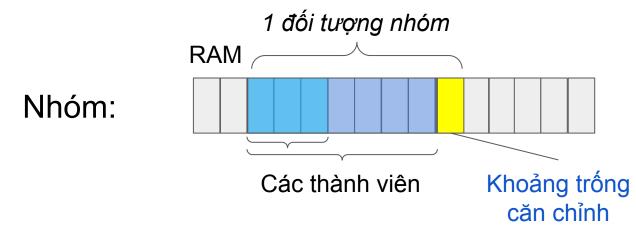
- Định dạng: (Kiểu-dữ-liệu){ /* danh sách khởi tạo */}
- Cấu trúc không tên:
 - \circ (struct point){.x = 100, .y = 200}
 - (const struct point){10.0, 11.5} Chỉ đọc
- Nhóm không tên:
 - (union gtype){.i = 1001}
 - (const union gtype){.d = 3.1415} Chỉ đọc
- Mảng không tên:
 - o (int []){1, 2, 3}

Các giá trị gộp có thể được lưu trong các đối tượng không tên, các đối tượng không tên chỉ đọc giống nhau có thể cùng là 1 đối tượng - được lưu trong 1 vùng nhớ.

Các mô hình bộ nhớ



Vùng nhớ của cấu trúc có thể có căn chỉnh ở giữa các thành viên hoặc ở cuối, nhưng không có căn chỉnh ở đầu.



Vùng nhớ của nhóm có thể có căn chỉnh ở cuối, nhưng không có ở đầu.

Hệ quả của mô hình bộ nhớ: Các kích thước

- Kích thước cấu trúc >= Tổng kích thước các thành viên
 - struct s1 { int x; char c; int y; };
 - sizeof(struct s1) >= sizeof(int) + sizeof(char) + sizeof(int);
 - struct s2 { float x; float y; float z; };
 - sizeof(struct s2) >= 3 * sizeof(float);
 - (> Nếu có căn chỉnh)
- Kích thước nhóm >= Kích thước thành viên lớn nhất
 - union g1 { float x; double d; };
 - sizeof(union g1) >= sizeof(double);
 - union g2 { struct xyz p; double d; };
 - sizeof(union g2) >= sizeof(struct xyz);
 - (> Nếu có căn chỉnh)

Ở 1 thời điểm 1 đối tượng nhóm thường chỉ lưu được giá trị của 1 thành viên - lưu vào thành viên nào thì đọc từ thành viên đó

Có thể kiểm tra khoảng trống dựa trên các kích thước và độ xê dịch của các thành viên tính từ địa chỉ của đối tượng chứa nó bằng macro offsetof (stddef.h)

Ví dụ 9.1. Kích thước của cấu trúc và nhóm

```
#include <stdio.h>
   #include <stddef.h>
 3
                                           sizeof(struct s1) = 12
   struct s1 { int x; char c; int y; };
                                           offset: 0 4 8
   struct s2 { int x, y, z; };
 5
                                           sizeof(struct s2) = 12
   union g1 { char s[9]; double d; };
                                           sizeof(union q1) = 16
   union q2 { float f; double d; };
 7
                                           offset: 0 0
 8
                                           sizeof(union q2) = 8
 9
   int main() {
10
     printf("sizeof(struct s1) = %zu\n", sizeof(struct s1));
     printf("offset: %zu %zu %zu\n", offsetof(struct s1, x),
11
              offsetof(struct s1, c), offsetof(struct s1, y));
12
     printf("sizeof(struct s2) = %zu\n", sizeof(struct s2));
13
14
     printf("sizeof(union q1) = %zu\n", sizeof(union q1));
     printf("offset: %zu %zu\n", offsetof(union q1, s),
15
16
              offsetof(union q1, d));
     printf("sizeof(union q2) = 2u\n", sizeof(union q2));
17
18
       Có thể nhận thấy thành phần căn chỉnh trong s1 (giữa c và y)
       và trong g1 (ở cuối)
```

Đặt tên kiểu

- Sơ lược cú pháp
 - typedef T D;
 - Nếu T D khai báo ident là 1 đối tượng thì typedef T D khai báo ident là kiểu của ident trong T D.
 - Sau đó có thể sử dụng tên kiểu để tạo đối tượng:
 - ident x;
 - x là 1 đối tượng có kiểu như ident trong T D;
- Ví dụ:
 - typedef struct point {double x, y; } point_s, *point_ptr;
 - point_s kiểu cấu trúc point, point_ptr kiểu con trỏ tới cấu trúc point;
 - point_s p2; // tương tự struct point {double x, y;} p2;
 - point_ptr pp; // struct point {double x, y;} *pp;
 - typedef union gtype { long I; double d; } gtype_u, *gtype_ptr;
 - gtype_u kiểu nhóm gtype, gtype_ptr kiểu con trỏ tới nhóm gtype;
 - gtype_u g2; // tương tự union gtype { long l; double d; } g2;
 - gtype_ptr pp; // union gtype { long I; double d; } *pp;

Mảng cấu trúc và nhóm

- Tương tự như đã làm với các kiểu đơn trị chúng ta cũng có thể tạo mảng của cấu trúc và mảng của nhóm:
 - (Đối chiếu và so sánh với struct point p;) struct point points[100]; // Mảng 100 cấu trúc point union gtype values[100]; // Mång 100 nhóm gtype struct point points2[] = {{100, 200}, {200, 300}}; // 1 chiều struct point triangles[][3] = { { {100, 100}, {200, 200}, {300, 300} }, }; // 2 chiều, kích thước 1 x 3 points[i]; // Phần tử thứ i trong points struct point *pp = points; // pp trỏ tới phần tử đầu tiên *(pp + i); // Phần tử thứ i trong points V.V..

Nội dung

- Khái niệm & cú pháp
- Các toán tử
- Xử lý cấu trúc và nhóm

Các toán tử thành viên

Định dạng:

- Toán tử chấm .: s.x
- Toán tử mũi tên ->: p->x

Điều kiện:

 Vế trái của toán tử. phải là đối tượng (có kiểu) cấu trúc hoặc đối tượng nhóm, vế trái của toán tử -> phải là con trỏ tới đối tượng cấu trúc hoặc con trỏ tới đối tượng nhóm.

Ý nghĩa:

- Toán tử. (chấm) được sử dụng để truy cập tới thành viên với tên như vế phải của cấu trúc hoặc nhóm ở vế trái.
- Toán tử -> (mũi tên) được sử dụng để truy cập tới thành viên với tên như vế phải của cấu trúc hoặc nhóm được trỏ tới bởi con trỏ ở vế trái.

Chuỗi toán tử thành viên được thực hiện từ trái sang phải

Ví dụ 9.2. Các thành viên

- struct point { double x, y; } p, *pp = &p;
 - o p.x, pp->x đều là thành viên x của p 1 biến kiểu double;
 - o p.y, pp->y đều là thành viên y của p 1 biến double khác.
- struct line { struct point p1, p2; int color; } I;
 - I.p1.x là thành viên x của thành viên p1 của l;
 - I.p2.y là thành viên y của thanh viên p2 của I.
- union gtype { char[8] s; struct point *pp;} g, *pg = &g;
 - o g.s, pg->s đều là thành viên s của g.
 - g.s[1], pg->s[1] đều là ký tự với chỉ số 1 của thành viên s.
 - g.pp->x, pg->pp->x đều là thành viên x của đối tượng được trỏ tới bởi thành viên pp của g.

Toán tử gán với kiểu cấu trúc và kiểu nhóm

Biểu thức gán có hiệu ứng giống như sao chép dữ liệu trong vùng nhớ của đối tượng ở vế phải vào vùng nhớ của đối tượng ở vế trái, ví dụ:

- struct xy s1, s2; /*...*/ s1 = s2;
 - Có thể tạo hiệu ứng tương tự với memcpy (string.h)
 - void *memcpy (void *dest, const void *src, size_t n);
 - memcpy(&s1, &s2, sizeof(struct xy));
- union xy u1, u2; /* ... */ u1 = u2;
 - memcpy(&u1, &u2, sizeof(union xy));
- Có thể gán các đối tượng có thành viên là mảng (nhưng không gán được mảng cho mảng):
 - struct arr10 { int elems[10]; } a1, a2;
 - a1 = a2; // memcpy(&a1, &a2, sizeof(struct arr10));
 - a1.elems = a2.elems; // Lỗi không gán được mảng cho mảng

Ví dụ 9.3. Biểu thức với nhiều toán tử

```
struct array {int n, *elems;} a, *pa = &a, b;
```

- &a.n; // Được hiểu là (&a).n hay &(a.n) ?
- a.elems[10]; // (a.elems)[10] hay a.(elems[10])?
- pa->elems[10]; // (pa->elems)[10] hay pa->(elems[10])?
- a == b; // Hợp lệ: a.n == b.n && a.elems == b.elems hay không hợp lệ?

Ví dụ 9.3. Biểu thức với nhiều toán tử₍₂₎

struct array {int n, *elems;} a, *pa = &a, b;

- &a.n; // Được hiểu là (&a).n hay &(a.n) ?
 - &(a.n) địa chỉ của thành viên n của a, bởi vì toán tử hậu tố
 có thứ tự ưu tiên cao hơn toán tử tiền tố.
- a.elems[10]; // (a.elems)[10] hay a.(elems[10])?
 - (a.elems)[10] tương đương (*(a.elems + 10)), bởi vì chuỗi toán tử hậu tố được thực hiện từ trái sang phải;
- pa->elems[10]; // (pa->elems)[10] hay pa->(elems[10])?
 - (pa->elems)[10]
- a == b; // Hợp lệ: a.n == b.n && a.elems == b.elems hay không hợp lệ?
 - Không hợp lệ: Toán tử so sánh == chỉ áp dụng được với 1 số kiểu định sẵn như đã học, người dùng tự định nghĩa khái niệm == cho các đối tượng thuộc kiểu tự định nghĩa.

So sánh con trỏ tới thành viên

Vó'i

```
struct xy { long x; double y; } s;
void *ps = &s, *px = &s.x, *py = &s.y;
các quan hệ so sánh sau có kết quả đúng:
    ps == px; // Không có khoảng trống trước phần tử đầu tiên
    px < py; // Vì y được khai báo sau x.</pre>
```

Với

```
union xy { long x; double y; } u;
void *pu = &u, *px = &u.x, *py = &u.y;
các quan hệ so sánh sau cho kết quả đúng:

px == py; // Vì các thành viên của nhóm có cùng địa chỉ
```

- a py -- pur // Không có khoảng trống ở đầu phám
- px == pu; // Không có khoảng trống ở đầu nhóm.

Các toán tử khác

Các kiểu cấu trúc và các kiểu nhóm được định nghĩa bởi người lập trình với ý nghĩa riêng của các giá trị. Vì vậy có nhiều toán tử không áp dụng được cho các kiểu do người dùng tự định nghĩa.

- Các toán tử sau không áp dụng được cho các đối tượng cấu trúc và các đối tượng nhóm:
 - Các toán tử đại số: Cộng (+), trừ (-), nhân (*), chia (/), các
 phép gán kết hợp, v.v..
 - Các toán tử so sánh: Lớn hơn (>), nhỏ hơn (<), bằng nhau
 (==) v.v.., và nhiều phép toán so sánh liên quan khác.
 - O V.V...

Có thể đóng gói các xử lý tương tự toán tử cho các đối tượng thuộc kiểu tự định nghĩa với hàm hoặc macro

Ép kiểu con trỏ

- Có thể ép kiểu con trỏ tới cấu trúc thành con trỏ tới phần tử đầu tiên của nó và ngược lại, ví dụ:
 - struct xy { double x, y; } s, *ps = &s;
 - double *px = (double*)ps; // Ok, px == &s.x;
 - ps = (struct xy*)px; // Ok, tmp == &s;
 - Có thể lấy con trỏ tới đối tượng từ con trỏ tới thành viên không phải thành viên đầu tiên dựa trên offsetof, ví dụ:
 - double *py = &s.y;
 - ps = (struct xy *)((char*)py offsetof(struct xy, y)); // OK, ps == &s;
- Đối với nhóm có thể ép kiểu con trỏ tới nhóm thành con trỏ tới thành viên bất kỳ của nó và ngược lại, ví dụ:
 - o union gtype {long I; double d; } g, *pg = &g;
 - long *pl = (long*)pg; // Tương đương với pl = &g.l;
 - double *pd = (double*)pg; // pd = &g.d;
 - pg = (union gtype *)pl; // OK pg == &g
 - pg = (union gtype *)pd; // Ok pg == &g

Thứ tự ưu tiên và chiều thực hiện chuỗi toán tử

Ưu tiên cao (thực hiện trước)

Tên toán tử	Ký hiệu	Ví dụ	Chuỗi
Chỉ số		a[9]	\Rightarrow
) Gọi hàm	()	f(3, 5)	
Thành viên	., ->	o.x, pp->x	
Tăng, giảm 1 (hậu tố)	++,	X++, X	
Tăng, giảm 1 (tiền tố)	++,	++x,x	<
Địa chỉ, truy cập	&, *	&x, *p	
Dấu (tiền tố)	+, -	+x, -x	
Đảo dãy bit, phủ định	~,!	~x, !x	
Dung lượng	sizeof	sizeof(int)	
Ép kiểu	()	(double)n	<
Nhân, Chia, Phần dư	*, /, %	x * y, x / y, x % y	⇒
Cộng, trừ	+, -	x + y, x - y	\Rightarrow
Dịch trái, dịch phải	<<, >>	x << y, x >> y	\Rightarrow
So sánh thứ tự	<, >, <=, >=	x < y	⇒
So sánh bằng	==, !=	x == y	\Rightarrow
AND theo bit	&	x & y	⇒
XOR theo bit	۸	x ^ y	\Rightarrow
OR theo bit		x y	\Rightarrow
AND lô-gic	&&	x && y	⇒
OR lô-gic		x y	⇒
Lựa chọn	?:	x? y: z	←
Gán đơn giản	=	x = y	<
Gán kết hợp	*=, /=, %=, +=, -=, <<=, >>=, &=, ^=, =	x *= y, x /= y,	26

Ưu tiên thấp (thực hiện sau)

Nội dung

- Khái niệm & cú pháp
- Các toán tử
- Xử lý cấu trúc và nhóm

Nhập, xuất với kiểu cấu trúc và kiểu nhóm

- Có thể nhập, xuất từng thành viên của cấu trúc, nhóm
 - (Có thể áp dụng đệ quy)
- Ví dụ nhập và xuất cấu trúc:
 - struct point {double x, y; } p;
 - scanf("%lf%lf", &p.x, &p.y); // Nhập các thành phần p.x, p.y
 - printf("x = %f, y = %f\n", p.x, p.y); // Xuất p.x và p.y
- Ví dụ tương tự với nhóm:
 - union gtype {long I; double d; } g;
 - scanf("%ld", &g.l); // Sử dụng g như 1 biến kiểu long
 - o printf("%ld", g.l);

Hàm, cấu trúc và nhóm

Hàm có thể nhận tham số là cấu trúc, nhóm và trả về cấu trúc, nhóm, qua đó có thể triển khai các thao tác phức tạp với cấu trúc và nhóm, trả về nhiều giá trị, v.v...

Ví dụ hàm với tham số và giá trị trả về là các cấu trúc:

Ví dụ tương tự với nhóm:

Ví dụ 9.4. Hàm với tham số con trỏ cấu trúc

Tính khoảng cách Euclide giữa 2 điểm trên mặt phẳng

```
#include <stdio.h>
 2 #include <math.h>
 3 struct point {
     double x, y; Con trỏ thường được sử dụng đế truyền đối tượng cấu
                   trúc cho hàm, đặc biệt là các cấu trúc có kích thước lớn.
  double distance(struct point *p1, struct point *p2)
      double x = (p1->x - p2->x), y = (p1->y - p2->y);
      return sqrt(x*x + y*y);
 8
 9
                                Nhập tọa độ 2 điểm: 1 1 3
   int main() {
10
11
       struct point p1, p2;
      printf("Nhập tọa độ 2 điểm: ");
12
       scanf("%lf%lf%lf%lf",
13
              &p1.x, &p1.y, &p2.x, &p2.y);
14
       printf("Khoáng cách = %.3lf\n",
15
              distance(&p1, &p2));
16
```

Ví dụ 9.5. Mảng cấu trúc

Nhập n điểm và tìm 1 điểm bất kỳ gần gốc tọa độ nhất

```
#include <stdio.h>
   #include <math.h>
   /* Định nghĩa point và distance */
   int main() {
   int n, j = 0;
   scanf("%d", &n);
     struct point a[N];
     for (int i = 0; i < n; ++i) {
        scanf("%lf%lf", &a[i].x, &a[i].y);
                                                 5
10
11
      struct point p0 = \{0, 0\};
12
      double min = distance(&p0, &a[0]);
      for (int i = 1; i < n; ++i) {
13
       double dist = distance(&p0, &a[i]);
14
15
        if (dist < min) {</pre>
                                                 0.5 6
16
          min = dist;
                                                 (1.000, 1.000)
17
          j = i;
18
19
     printf("(%.31f, %.31f)\n", a[j].x, a[j].y);
20
21
```

Ví dụ 9.6. Sử dụng nhóm như 1 tập kiểu

```
#include <stdio.h> gtype đại diện cho 1 tập kiểu
   typedef union generic type
 3
     long 1;
    double d;
 4
   } qtype;
   gtype sum d(gtype g1, gtype g2) {
      return (qtype) \{ .d = q1.d + q2.d \};
 7
8
   qtype sum l(qtype q1, qtype q2) {
      return (qtype) \{ .1 = q1.1 + q2.1 \};
10
11
    int main() {
                                       Nhâp vào 2 số thưc: 1.5 3.7
12
13
     qtype q1, q2;
                                       1.500 + 3.700 = 5.200
14
     printf("Nhâp vào 2 số thưc: ");
                                       Nhâp vào 2 số nguyên: 10 20
15
    scanf("%lf%lf", &q1.d, &q2.d);
                                       10 + 20 = 30
    printf("\%.3f + \%.3f = \%.3f\n",
16
17
             q1.d, q2.d, sum d(q1, q2).d);
     printf("Nhập vào 2 số nguyên: ");
18
19
     scanf("%ld%ld", &q1.1, &q2.1);
     printf("%ld + %ld = %ld\n",
20
21
             q1.1, q2.1, sum l(q1, q2).1);
22
```

Ví dụ 9.7. Thêm thông tin kiểu cho đối tượng

Thử nghiệm giản lược định dạng hàm xuất dữ liệu

```
#include <stdio.h>
   typedef union generic type {
     long 1;
  double d;
  } gtype;
 5
   typedef struct tagged gtype {
     char tag;
     qtype val;
 8
    } tqt;
 9
   void print tgt(tgt x) {
10
      if (x.taq == 'l') {
11
12
       printf("%ld: long\n", x.val.l);
     } else if (x.taq == 'd') {
13
       printf("%g: double\n", x.val.d);
14
15
16
17
    int main() {
                                           10: long
      tgt v1 = {'l', {.l = 10}},
18
                                           3.5: double
          v2 = \{'d', \{.d = 3.5\}\};
19
20
  print tqt(v1);
21
     print tqt(v2);
22
```

